An Evaluation of Diarrhoea and Enteric Infection Surveillance Methods in Urban Informal Settlements and Refugee Camps

Ryan Trevor Titus Rego

Division of Health Sciences

Warwick Medical School

University of Warwick

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Health Sciences

May 2021

Table of Contents

List of Illustrations

List of Tables

Dedication

I dedicate this thesis to my parents, Derek and Cherryl, and my grandparents, Angela, Bertram, Hilda, and Titus, for always supporting me and pushing me in life.

Acknowledgements

I would firstly like to say thank you to my supervisors: Prof Richard Lilford, Dr Sam Watson, and Prof Paramjit Gill. Without their mentorship, (lots of) patience, and kindness this PhD would not have been possible, let alone to the level they pushed me to.

I would also like to thank everybody else at Warwick and Birmingham that supported me in this thesis, whether it was in applying for funding, booking flights, purchasing supplies, or any of the other tasks I was constantly asking for help with: Jo Sartori, Simon Smith, Ria Wilson, Melita Harris, and Tom Lingard.

I'd also like to thank my colleagues and friends in Bangladesh for their part in designing our studies in Bangladesh, implementing the field work, running the lab tests, and now in publishing our results: Dr Sirajul Islam, Dr Mohammed Yunus, Dr Syed Asif Abudllah, and Atique Ul-Alam.

Finally, I'd like to thank my family and friends both in the UK and abroad for supporting me during my PhD, particularly my parents, Derek and Cherryl, and my partner, Lilly.

Funding

I thank the University of Warwick for providing my studentship to pursue my PhD. I thank the National Institutes for Health Research Global Health Research Unit on Improving Health in Slums (16/136/87) for providing funding for the travel needed for my PhD research. Research costs for Chapters 3 and 4 were also supported by the University of Warwick through internal funding.

The views expressed in this thesis are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care. The funder had no role in study design, implementation, or interpretation.

Declaration

This thesis is submitted to the University of Warwick in support of my application to the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree.

Parts of this thesis have been published by the author:

Chapter 3: Rego, R., Watson, S., Ishengoma, P., Langat, P., Otieno, HP., Lilford, R.. Effectiveness of SMS messaging for diarrhoea measurement: a factorial cross-over randomised controlled trial. BMC Med Res Methodol 20, 174 (2020). https://doi.org/10.1186/s12874-020-01062-3

Parts of this thesis have been submitted for publication by the author, and may be available in pre-print:

Chapter 4: Rego, R., Watson, S., Lilford, R., Gill, P. The Impact of Diarrhoea Measurement Methods for Under-Fives in Low and Middle Income Countries on Reported Diarrhoea Rates: A Systematic Review and Meta-Analysis of Methodological and Primary Empirical Studies.

Ryan Trevor Titus Rego

The First of May 2021

Abbreviations

ELISA: Enzyme-linked Immunosorbent Assay

EPEC: Enteropathogenic Escherichia Coli

ETEC: Enterotoxigenic Escherichia Coli

EWARS: Early Warning and Response System

FDMN: Forcibly Displaced Myanmar National

ICDDR,B: International Centre for Diarrhoeal Disease Research, Bangladesh

ISRCTN: International Standard Randomised Controlled Trial Number

LMIC: Low and Middle Income Countries

NGO: Non Governmental Organization

PCR: Polymerase Chain Reaction

RDT: Rapid Diagnostic Test

SIR: Sustainable, Infected, Recovered

SMS: Short Message Service

STEC: Shiga toxin-producing Escherichia Coli

UN: United Nations

UNDP: United Nations Development Program

UNHCR: United Nations High Commissioner for Refugees

WASH: Water, Sanitation, and Hygiene

WHO: World Health Organization

Abstract

Surveillance of diarrhoea and enteric infection among infants is key to improving infant health in LMICs. However, the performance of current surveillance methods in estimating rates of diarrhoea and enteric infection have been called into question by various academics. Diarrhoea rates are often estimated through subjective measures such as community surveys. These measures may introduce variation into the estimated rates. Rates of enteric infection are often estimated using diarrhoea as a proxy marker. However, it is possible that diarrhoea does not classify enteric infection well. In this thesis, I examine variation in estimates of diarrhoea rates associated with the method used; and the extent of misclassification when using diarrhoea as a proxy marker of enteric infection.

To examine variation in estimates of diarrhoea rates associated with the method used, I conducted two studies: 1) a systematic review and meta-analysis of the literature; and 2) a randomised controlled trial (RCT) in a Tanzanian urban informal settlement. My findings in both confirmed that the methods used in estimating diarrhoea rates are associated with variations in estimates. The RCT in Tanzania also showed that text message surveys are a suitable substitute for in-person surveys.

To examine the reliability of diarrhoea to classify enteric infection, I conducted a 'diagnostic' study in a Bangladeshi Refugee Camp. The evidence showed that diarrhoea misclassified enteric infection, offering no better a classification of enteric infection than flipping a coin. I also found evidence of clear seasonality in infection.

I have demonstrated that the methods used in diarrhoea surveillance can influence estimated rates, and that diarrhoea widely misclassifies enteric infection when used as a proxy marker. For this reason, I hazard against the use of the current methods in epidemiological surveys when trying to estimate rates of diarrhoea and classify cases of enteric infection. Alternative, less subjective, methods, such as stool sampling for enteric infection, should be used.

1. Introduction

1.1. Background

Diarrhoea is the second highest cause of under-five death after the first month of life in low and middle income countries (LMICs)[1]. In 2017, diarrhoea claimed 424,000 under-five lives[1, 2]. Diarrhoea additionally puts children at an increased risk of stunting, contributing to the 30% prevalence of stunting among under-fives in LMICs. Diarrhoea may also contribute to impaired cognitive development, possibly attributable to the 200 million children globally with impaired cognitive development[3-5].

Under-fives living in informal settlements with high population densities, such as refugee camps and slums, are at an increased risk of diarrhoea compared to their counterparts in rural areas and urban formal settlements[6]. This risk is due to factors that increase the transmission of enteric pathogens, including poor water, sanitation, and hygiene (WASH) infrastructure, proclivity to natural disasters, and high population densities[7, 8]. As a result of the increased risk of diarrhoea (as well as other contributing reasons, such as malnutrition and other infections), under-fives in urban informal settlements have a 34% higher chance of dying before their fifth birthday than their peers in formal or rural areas, as well as a 14% higher chance of being stunted[9, 10]. In order to lower the risk of diarrhoea, WASH interventions are often deployed to reduce enteric infections, the main cause of diarrhoea. WASH interventions range from simple, such as the construction of latrines in refugee camps, to large, such as the introduction of sewage and piped water systems. A large amount of funding and workforce capacity goes into the implementation and evaluation of these systems every year.

Currently, 900 million people live in urban informal settlements globally, with a further 2.6 million people living in formal refugee camps (many more refugees live in informal camps and urban settings)[11, 12]. The number of people living in urban informal

settlements has been steadily increasing, and is expected to double or even triple by 2050[13]. Additionally, given ongoing conflicts and natural disasters, the number of people living in refugee camps is also expected to grow. As such, attention must be paid to the abatement of diarrhoea in urban informal settlements and refugee camps, particularly through WASH interventions.

The Importance of Proper Surveillance

To reduce the burden of under-five diarrhoea, particularly pathogenic diarrhoea caused by enteric infection, proper surveillance systems are needed to estimate rates of diarrhoea and enteric infection. Data provided by surveillance systems have several uses, but the most important are: 1) aiding public health authorities in detecting and responding to outbreaks; and 2) evaluating WASH interventions designed to limit the spread of enteric infection.

Outbreak Detection

Throughout LMICs, outbreaks of pathogenic diarrhoea (diarrhoea caused by enteric infections) are at the forefront of public health concerns. Pathogenic diarrhoea outbreaks are caused by increased transmission of enteric pathogens. There are various factors that contribute to outbreaks, including the emergence of new strains, seasonality, or natural disasters. Often, these outbreaks take place in areas with already high rates of enteric pathogen transmission that have suffered an event that exacerbate the underlying issues. A recent example is the 2010 Haitian Cholera Outbreak. Caused by hurricanes and floods, the outbreak claimed 8,183 lives[14]. Numerous other examples also exist for other enteric pathogens, such as Rotavirus, Norovirus, and Hepatitis E. In order to properly detect and control these outbreaks, routine surveillance of these pathogens is needed.

During endemic spread of pathogens (when an outbreak is not taking place), routine health surveillance systems feed data on disease levels back to public health authorities. The data allow health authorities to track incidence rates of infection. When incidence rates reach a pre-defined outbreak threshold, health authorities declare that an outbreak has occurred. The threshold rates differ by pathogen type. For example, the WHO defines a cholera outbreak as a single case (with evidence of community transmission)[15]. However, for many other enteric pathogens which are more endemic than Cholera, incidences at 1.5 times the baseline of the previous three weeks are often used as the treshold[16]. The declaration of an outbreak leads to the provision of enhanced interventions, such as water chlorination and health education, and often the provision of extra funding. The effectiveness of surveillance systems to declare outbreaks relies on the provision of timely, accurate, and useful data.

A key challenge to controlling an outbreak is determining its source. Proper epidemiological data and analysis is the backbone for this. John Snow famously, for example, used rudimentary geospatial techniques through mapping cholera outbreak data in 1854 London to determine that a single water pump was the source of disease spread, allowing for clear actions to abate spread (in this case removing the pump handle)[17]. With these data, policy and program makers can properly take targeted steps to stop the outbreak, whether that be (as in the case of Snow) cutting of a contaminated water source, or putting in place temporary, enhanced water, sanitation, and hygiene systems.

Evaluation of Interventions

Around the world, interventions to reduce the burden of pathogenic diarrhoea have been deployed by governments, communities, and non-governmental organizations (NGOs). The most common of these interventions are WASH interventions. WASH

interventions aim to reduce the burden of pathogenic diarrhoea by reducing the spread of enteric pathogens. They do this by forming a barrier between enteric pathogens in the environment and susceptible people. These range from the simple, such as the provision of hand soap and chlorinated water, to the complex, such as the construction of large sewage systems. In addition to WASH interventions, health interventions, such as Zinc supplementation and vaccination, and nutritional interventions, such as the provision of enriched food, are used with the aim of reducing the burden of enteric infection and ultimately diarrhoea (both pathogenic and non-pathogenic)[18].

Every intervention requires proper data on population-level enteric infection rates to properly evaluate it. Evaluations allow for comparison against both the status quo and against competing interventions, resulting in data that allows program makers to choose the best, most appropriate intervention for the context. However, recent reviews of the effectiveness of various WASH interventions have revealed that there is little evidence for the efficacy of common WASH interventions[18]. The lack of evidence on effectiveness was also seen in several evaluations, such as randomised control trials, prospective trials, and retrospective evaluations. Most of these studies used carer-reported diarrhoea as their main endpoint. Further, two of the three recent integrated large WASH trials, the WASH Benefits trials from Bangladesh and Kenya, and the SHINE trial, have shown null results[19]. These studies also used carer reported diarrhoea as their main endpoint. In their recent review, Wolf et al. (2018) discuss that this may not be due to intervention ineffectiveness, but due to issues in the methods used to estimate rates of diarrhoea and enteric infection[18, 20]. No literature exists on the extent of these issues. As such, further research is needed on the impact of estimation methodology on reported diarrhoea rates, as well as the extent of misclassification when using diarrhoea to act as a proxy marker of enteric infection.

1.2.The Unknown Impact of Diarrhoea Estimation Methods on Reported Diarrhoea Rates

An Epidemiological Model of Diarrhoea Spread

To explain the different methods of diarrhoea and enteric infection surveillance, I will begin by looking at how different types of diarrhoea and enteric infections spread in populations, and how people may be affected by them. I will do this by creating an epidemiological model of diarrhoea spread, presented in a modified susceptible, infected, recovered (SIR) model (Fig. 1). All people start as 'Healthy', without any diarrhoea or infection. Once a healthy person becomes ill, they can either have an enteric infection (sometimes leading to pathogenic diarrhoea), or non-pathogenic diarrhoea, caused by things such as stress, nutritional intolerances, or non-enteric infections[21, 22]. Both enteric infection and non-pathogenic diarrhoea can exist in symptomatic states where a person has not sought treatment, or symptomatic states where a person has sought treatment. These may eventually lead to death (it is worth noting however that there is no literature on mortality due to non-pathogenic diarrhoea). Enteric infection exists in related states: asymptomatic infection, which may or may not progress to symptomatic infection; recovered with carriage, where a person no longer has symptoms but is still infectious; or recovered with immunity, in which the previously infected person is no longer infected but has some immunity. A person who is healthy and is then vaccinated will also flow into the 'recovered with immunity' state.

It is also important to note how WASH interventions fit into this diagram. In a very simplistic manner, WASH interventions prevent a healthy person from becoming infected. They do this by forming a barrier between infected people and the environment (e.g. through faecal sludge management); pathogens in the environment

and healthy people (e.g. through safe water); and between infected people and healthy people (e.g. hand washing). This is represented by the 'WASH Barrier' on the diagram.

There are several methods of estimating diarrhoea rates in populations, each of which estimates rates in people with different disease states. Mortality surveillance, for example, estimates rates from those who have died from diarrhoea[23]. Passive surveillance estimates rates from hospitalised cases of diarrhoea[23]. Active surveillance, through door to door surveying, estimates rates from those who have diarrhoea in communities[23]. However, none of these methods are able to discriminate between pathogenic and non-pathogenic diarrhoea, nor are able to estimate rates from all those with an enteric infection regardless of disease state (those shaded green on the diagram). As described above, we are most interested in those who have an infection – regardless of disease state. Stool sampling for molecular surveillance, the classification of infection through the collection of stool samples, is able to make this distinction. In addition, other methods such as environmental sampling may be used[24]. It is possible that the use of differing surveying methods which estimate rates from different disease states may result in bias (systematic misreporting in a certain direction) and error (random misreporting in no certain direction) attributable to the estimation method used.

Figure 1 SIR Model of Pathogenic and Non-Pathogenic Diarrhoea Spread in an Environment

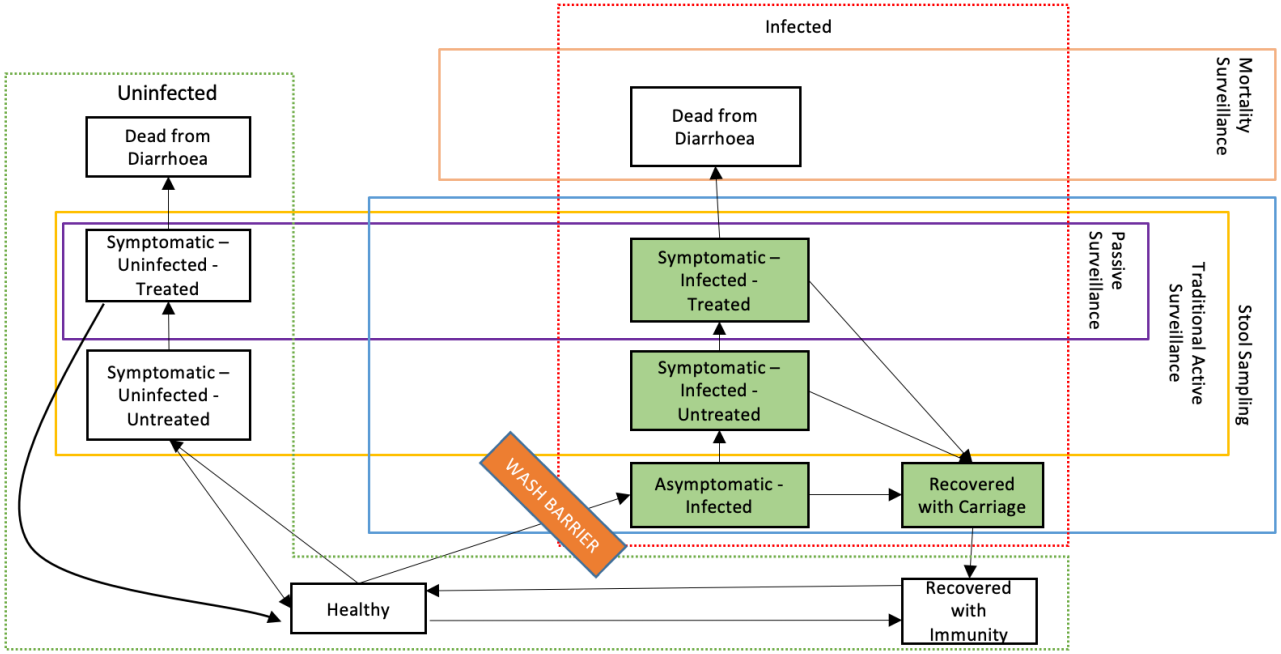


Figure 1 Caption: A modified Susceptible, Infected, Recovered (SIR) model of pathogenic and non-pathogenic spread in an environment with overlays for surveillance types. All boxes are states a person can lie in other than the orange box which is the barrier formed by WASH interventions. Green states are those with an active enteric infection, which WASH interventions can impact.

Variability of Estimated Diarrhoea Rates Based on Method Used

The first concern in the estimation of diarrhoea rates is variation attributable to the method used to estimate rates. As mentioned above, the two most common methods of population-level diarrhoea estimation are passive and active surveillance. Passive surveillance, the estimation of rates from hospital cases for diarrhoea, is found commonly in routine surveillance for endemic diseases. Notable examples include the World Health Organization's Early Warning and Response System (EWARS) and Public Health England's Notifiable Disease and Causative Organisms Reporting System (NDCORS). During humanitarian emergencies, EWARS collects data from health facilities for visits of communicable diseases, such as diarrhoea or chickenpox[25]. In the UK, NDCORS collects similar data from public health laboratories, in addition to health facilities[26]. Active surveillance, door to door questioning of diarrhoea, is found commonly during evaluations of WASH interventions[23].

Possible Variations Associated with Passive Surveillance

As not all people with diarrhoea visit health care facilities, passive surveillance may be associated with lower estimated population-level diarrhoea rates than active surveillance. There are two main reasons why a child with diarrhoea may not visit a health care facility: 1) children generally do not visit health facilities unless their diarrhoea is severe, only 2% of cases[27, 28]; and 2) some populations may face issues accessing healthcare due to physical distance, legal barriers, fear, or affordability. Marginalized groups, such as refugees and those living in informal settlements, particularly face these barriers[29-31]. Further, even when accessing healthcare, carers may take their children to healthcare providers that do not report to the area's passive surveillance system – such as traditional healers and unregistered providers.

Even when children with diarrhoea visit the appropriate health facility, health care providers may incorrectly diagnose the disease, or may diagnose the disease with differing case definitions[32]. While documented for other diseases, this has not been empirically demonstrated for diarrhoea. Further, health facilities may fail to report cases – particularly on weekends and during holidays.

Possible Variations Associated with Active Surveillance

Active surveillance systems can differ in quantitative means, such as the recall period of questioning and the frequency of questioning, as well as qualitative means, such as influence by enumerators or by respondents experiencing interventions. Given the subjectivity of active surveillance methods, these can all introduce variations in reported diarrhoea rates.

Previous work has indicated that the length of recall over which carers are asked to remember diarrhoea may impact reported diarrhoea rates. This variation attributable to recall length is due to recall bias, when carers forget events which have happened in the past. Evidence indicates that shorter recall periods yield higher, and likely more accurate, reported diarrhoea rates than longer periods. In rural Kenya, Feikin et al. (2010) saw that estimated diarrhoea prevalence dropped from 18% with two days of recall to five percent with eleven days of recall. Evidence also indicated that recall bias is more impactful for moderate diarrhoea than severe diarrhoea[33]. Zafar et al. (2010) found that severe diarrhoea was twice as likely to be reported as moderate diarrhoea during periods of long recall (one week compared to two days)[34].

In addition to carers forgetting if their child had diarrhoea, carers may also be swayed by how often, and the manner in which, surveyors question them on diarrhoea. Bias and error introduced by the frequency of surveying is known as respondent fatigue. While not examined yet empirically, it has been hypothesized

that repeated surveying at short frequencies can cause a decrease in reliability due to carers becoming tired of asking questions[35]. Reactivity, when participants answer questions differently due to knowing that they are being observed, is of additional concern. Reactivity has not, however, been empirically examined for diarrhoea estimation[36]. Further, parents may not even be able to properly identify what is and is not diarrhoea[37].

Misclassification when using Diarrhoea as a Proxy Marker of Diarrhoea

As discussed previously, the two main uses of diarrhoea surveillance, outbreak detection and WASH intervention evaluation, aim to estimate rates of enteric infection. The unwritten assumption is that diarrhoea is a proxy marker of enteric infection. There is reason to believe that diarrhoea may misclassify infection when used as a proxy marker, due to asymptomatic infection and non-pathogenic diarrhoea. While some may argue that diarrhoea is the main clinical outcome of interest, and that one should not be concerned about misclassification, I disagree. Asymptomatic infection is important as it measures the general population risk of infection, which is both important to outbreak surveillance and can be impacted by WASH trials. Those with asymptomatic infections may also pass pathogens on to another person who may experience significant morbidity and mortality. Non-pathogenic diarrhoea is also not of interest in this situation as it will not result in outbreaks, and cannot be controlled by WASH interventions. It is important to note, however, that I say this only for epidemiological surveys and not for a clinical setting.

Okitsu et al. (2020) recently estimated that in Matlab, a rural area of Bangladesh, 80% of children who did not report any diarrhoea had an enteric infection[38]. George et al. (2018) found similar results, reporting that 76% of children with enteric infections in another area of rural Bangladesh reported no diarrhoea[39]. With this large rate of asymptomatic cases it is likely that asymptomatic spread precedes any

spike in the number of symptomatic cases. As such, estimation using only symptomatic cases precludes early detection of outbreaks and the deployment of resources to the most at need areas, as well as possibly offering false assurances of WASH intervention effectiveness. Diarrhoea can also be non-pathogenic – caused by maladies such as non-enteric infections (such as HIV, which sometimes causes HIV enteropathy), poor nutrition, and stress[21, 40]. However, the rates of non-pathogenic diarrhoea have not been examined. Further, these issues of asymptomatic infection and non-pathogenic diarrhoea have not been examined together to determine the extent and consequences of misclassification when using diarrhoea as a proxy marker of infection.

1.3. Unexamined Alternative Methods for Diarrhoea and Enteric Infection Surveillance

I have above described some of the issues in estimating diarrhoea rates using active and passive surveillance, and in using diarrhoea as a proxy marker of enteric infection. Below I discuss possible methods to improve these.

<u>Methods to decrease variability when estimating diarrhoea rates</u>

SMS Surveys to Reduce Cost and Reactivity

In-person active surveillance of diarrhoea is highly cost and resource intensive, and can possibly introduce reactivity into estimates. With 90% of people living in LMICs having access to mobile phones, a possible solution to this is questioning using text messaging[41, 42]. During an evaluation of Family Planning programs in Malawi, Pattnaik et al. (2020) found that mobile phone interviews obtained similar results to in-person surveys, while drastically cutting down on cost[43]. Additionally, during the 2014-2015 Liberian Ebola Outbreak, Kuehne et al. (2016) found that SMS surveying of random phone numbers to obtain information on Ebola morbidity, mortality, and care seeking obtained a response rate of 15%[44]. Similarly, L'Engle et al. (2017) obtained a 31% response rate in an SMS survey on health and

demographics in Ghana[45]. However, L'Engle et al. additionally found that younger, urban, more educated, and male respondents were more likely to reply to the survey, which may introduce further variation in estimated rates through selection bias[45].

Pictorial Surveys and Visual Stool Analysis to Reduce Perception Bias

Perception bias, bias resulting from the inability of carers to correctly identify if a stool is diarrhoea or not, can possibly be remedied through the use of pictorial surveys – showing carers pictures of stool and asking them to identify which stools their child has had, rather than verbally asking a carer if their child has had diarrhoea[46]. Past evaluations of pictorial surveys have used stool scales designed for adults (such as the Bristol stool scale) and have been hospital based[47]. As an alternative to the Bristol stool scale, the Amsterdam stool scale has been successful in clinical settings for stool classification among infants – but has not been tested for community based diarrhoea surveys. If successful in reducing perception bias, they offer a useful tool for community based diarrhoea surveillance.

Another possible method to reduce the impact of perception bias in diarrhoea surveillance is to eliminate carer reporting all together. This can be done through the collection of a stool sample from the infant which is then inspected by a trained researcher or health care provider. Voskujil et al. (2017) used this method successfully in hospital based surveillance, but it has not been examined in the community[37]. This method, however, does have various drawbacks – including needing to revisit households to collect stool, resources required to collect stool, inability to measure diarrhoea which may have occurred in the past (thus estimating a point prevalence and not longitudinal prevalence), and training of staff.

Alternative Methods for Classifying Enteric Infection

The most common method of detecting pathogens in stool is through polymerase chain reaction (PCR), a laboratory method which can detect pathogenic DNA in stool and identify the type of pathogen. This method is also capable of determining the amount of pathogen in stool, through this is of higher cost and more resource-intensive. PCR stool testing has previously been used to determine prevalent causes of disease in clinical settings, and more recently to estimate the rates of asymptomatic carriage in populations, but is still seldom used in epidemiological surveys estimate infection rates throughout communities[38, 48]. If there is indeed a disconnect between reported diarrhoea and infection, stool testing may provide a way of reducing this issue. Further, stool testing, compared to disease surveillance, allows for discrimination by pathogen type. As different pathogens may be sensitive to different control measures, this can aid greatly in public health program and policy. PCR, however, is an expensive technique, and requires complex lab equipment with highly trained staff – creating large barriers to testing in emergencies, rural settings, and low income settings. As such, alternatives are needed to facilitate widespread rollout.

Protein Testing

Protein biomarkers may be a low-cost but more accurate proxy for infection (when compared to carer-reported diarrhoea). There are several possible proteins that can fit this use, but calprotectin has been shown to be the most useful, particularly as a proxy for bacterial infection. Shastri et al. (2008) found that having elevated calprotectin levels classified bacterial infection with an 83% sensitivity and 37% specificity in adults[49]. Chen et al. (2012) similarly found that those with bacterial infections had elevated calprotectin levels when compared to those with viral infections, and also found a positive association between calprotectin levels and the severity of disease (as measured by the Vesikari score)[50]. Chen et al. (2011)

obtained similar results for lactoferrin, another protein associated with bacterial infection. Neither, however, was a sensitive marker for viral diarrhoea[51].

1.4.Thesis Justification and Structure

This thesis aims to explore some of the issues discussed above, with the aim of evaluating the variability of current diarrhoea surveillance methods in estimating diarrhoea rates, and the extent of misclassification when using diarrhoea as a proxy marker of enteric infection.

I begin by looking into issues in the variability of current diarrhoea surveillance methods in estimating diarrhoea rates. As discussed above, past research has hypothesized that passive surveillance methods may result in lower estimated diarrhoea rates than the actual rates, and factors specific to active surveillance may cause bias and error in estimated rates. Chapter Two "The Impact of Diarrhoea Estimation Methods for Under-Fives in LMICs on Reported Epidemiology: A Systematic Review and Meta-Analysis of Methodological and Primary Empirical Studies" evaluates this by systematically reviewing all studies between 2000 and 2018 that estimate diarrhoea rates among under-fives in LMICs, and meta-analysing their results. This chapter aims to determine the frequency of different diarrhoea estimation methods, and the impact that each methodology has on estimated diarrhoea rates. The study also looks for evidence of reactivity.

In Chapter Three, I then investigate SMS messaging as a method for reducing the variability introduced by active surveillance methods. This chapter presents a factorial randomized control trial from Mwanza, Tanzania, which evaluates the use of SMS messaging for diarrhoea rate estimation in urban informal settlements. The chapter also presents evidence on the impact of recall period, questioning frequency, question load, and incentivization on survey response rates and reported diarrhoea rates.

The second half of the thesis addresses the aforementioned issues in using diarrhoea as a proxy marker of enteric infection. Chapter Four "A Comparison of Traditional Diarrhoea Measurement Methods with Microbiological and Biochemical Indicators in the Cox's Bazar Displaced Persons Camp: A Cross-Sectional Observational Study" is a cross-sectional study from the Cox's Bazar refugee camp. In this chapter I evaluate the extent of misclassification when using carer-reported diarrhoea as a proxy marker of enteric infection. The chapter also compares the standard method of carer reported diarrhoea against several alternatives: pictorial questioning, use of proxies, visual stool examination, and presence of protein markers of infection.

I conclude by showcasing the utility of the direct measurement of enteric infection, rather than diarrhoea, in epidemiological studies and surveillance activities. Chapter Five "Prevalence, Determinants, and Spatial Variance of Enteric Infection by 27 Endemic Pathogens in the Cox's Bazar Forcibly Displaced Myanmar Nationals Camp: A Cross-Sectional Observational Study" presents the epidemiology and risk factors of enteric infection in the Cox's Bazar Forcible Displaced Myanmar Nationals Camp.

Chapter Six will discuss the implications of the four studies included in this thesis, specifically looking at how estimation methods may impact reported disease rates, the association between disease and infection, and the various uses of disease and infection epidemiology.

1.5.An Overview of the Study Settings

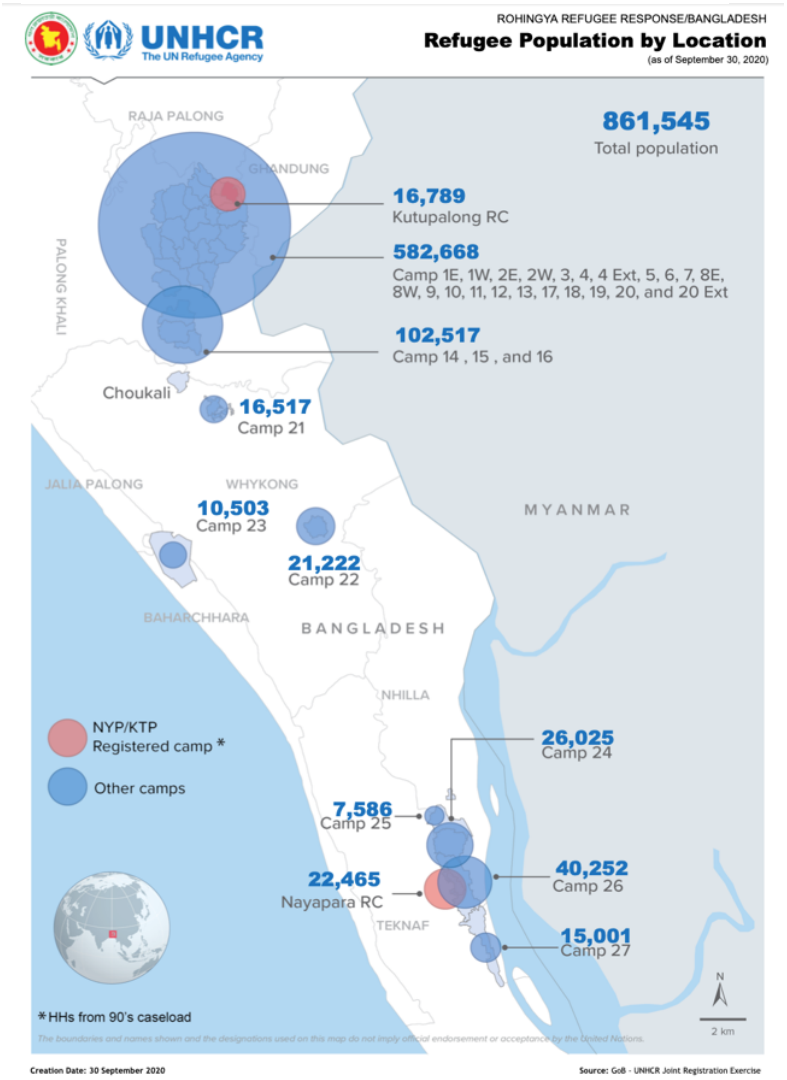
Cox's Bazar Forcibly Displaced Myanmar Nationals (FDMN) Camps

The Cox's Bazar FDMN camps are a collection of 29 individual FDMN camps in the Cox's Bazar region of Bangladesh. Cox's Bazar is a coastal region on the Bay of Bengal, and borders Burma to the east, separated by the narrow Naf river. As of

March 2019, the camps were officially home to 909,000 FDMNs (Fig. 2)[52]. However, the actual number is likely to be much higher given a lack of formal registration of the FDMNs. The camp area is very dense, with almost no separation between homes (Fig. 3). Below I present the Political, Economic, Social, Technological, and Environmental factors affecting this group.

- Political: The FDMNs living in Cox's Bazar, primarily members of the Rohingya ethnic minority group in Burma's Rakhine state, sought refuge from a violent genocide in Burma which has been raging in waves since the 1962 Military Coup[52]. Since this time, millions of Rohingya have crossed the Naf River into Bangladesh. The most recent influx between 2016 and 2017 saw close to one million people flee to Cox's Bazar[52]. While some of the FDMNs use Cox's Bazar only as an intermediate point before making onward journeys to countries such as Malaysia and Thailand, the majority remain in Cox's Bazar. The Bangladeshi government permits the FDMNs to stay within the boundaries of Cox's Bazar, but does not recognize the FDMNs as refugees, circumventing the protections afforded under the 1951 convention and the 1967 protocol on refugees. The FDMNs are generally not able to leave the camp area.
- Economic: Close to 80% of those living in the FDMN camps report no household income, with 35.1% living in debt[53]. While the FDMNs receive in kind supplies of food, cleaning and cooking supplies, and other necessities, income is needed to supplement these to buy items such as spices, meat, clothes, or shelter items. Of the FDMNs who do work, they are restricted to working in the camp. A majority of those working seek work with local NGOs on construction projects, with the more affluent FDMNs owning small shops in the camp. Women are also occasionally economically active, often working in tailoring.

- Social: The FDMNs living in Cox's Bazar are Muslims from Burma. Being Muslim, to some degree, facilitates acceptance by the Bangladeshi host community. There is, however, still a large degree of animosity between the FDMNs and local host community, often resulting in violence. Gender based violence is also a concern in the camps, and is perpetuated by the FDMNs themselves, the host community, and aid workers. The FDMNs in general have a low level of education[53]. Children in the FDMN camps are able to attend schools provided by the NGOs, but are not allowed to learn Bangla. FDMN communities and families are strong patriarchal in nature, with communities led by a chosen male official, often a religious or political leader from Burma, known as a Maji.
- Technological: Technology is scarce in the FDMN camps in Cox's Bazar, owing to a lack of infrastructure (including electricity), legal barriers, and poverty. Many houses use solar electricity to power lighting and fans, but beyond that generators must be used for larger appliances – a prohibitive cost to most. Most FDMN households have access to a mobile phone, despite mobile phones being banned in the FDMN camp.
- Environmental: The refugee camps in Cox's Bazar are located in and around a forest region, near the Naf River and Bay of Bengal. This area has very steep topography, leading to widespread landslides and flooding. Further, due to rivers in and around the camps, drowning is an issue, particularly for children. In general, population density is very high in the camps.



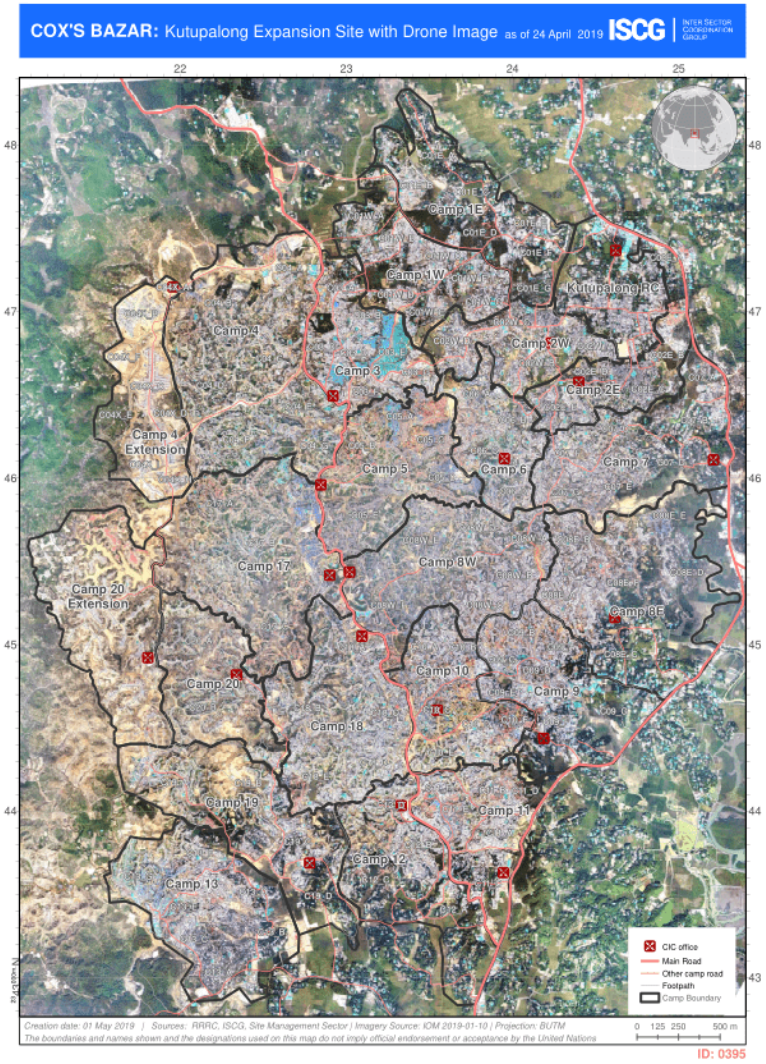


Figure 2 (Left): Location and Areas of all FDMN camps in Cox's Bazar, with population sizes. My study took place in Camp 24, Leda. From UNHCR, 2020

Figure 3 (Right): An aerial view of the Kutupalong camp, the largest camp area in Cox's Bazar. From ICSG, 2019.

Mwanza, Tanzania

Mwanza, Tanzania is a large city on the banks of Lake Victoria. The city is home to 2.8 million people, of which 18% are under the age of five. My research in Mwanza focused on three informal settlements: Kilimanhae, Igogo, and Unguja (Fig. 4).

- Political: Tanzania is a democratic nation with the ruling party (Chama Cha Mapinduzi) being center-left. This results in relatively liberal policies, and wide-spread government services and welfare. These policies have been cited as a contributing factor to the recent investments in sanitation infrastructure in Tanzania's urban centers, particularly Dar es Salaam, Arusha, and Mwanza[54]. At a local level, each of the informal settlements I worked in had their own local leader, who was appointed by the ruling party in Tanzania in Mwanza. There is generally no opposition to these appointments, as the local leaders are, for the most part, respected elders in their communities.
- Economic: Tanzania's GDP per capita was 3,222USD (adjusted for purchasing power parity, PPP) in 2018, with Mwanza having a slightly higher GDP per capita of 3,289USD (PPP)[55]. In 2017, Tanzania's GINI index was 0.40, lower than the global average of 0.63, indicating that wealth is near evenly distributed in Tanzania[56]. However, Tanzania has a high poverty rate – 26.4% in 2018[56]. That being said, great strides have been made in addressing the poverty rate, with consistent year on year decreases in the country's poverty rate[56].
- Social: While no city-specific figures could be found, urban areas in Tanzania (such as Mwanza) have religiously mixed populations, with a majority being Christian and a minority being Muslim[57]. This does sometimes result in conflict between groups[57]. Mwanza itself, similarly to Tanzania as a whole, has a young population – with 46.7% of the population under the age of

14[58]. When examining markers of gender equity, 37% of married women were married before the age of 18, and 29% of females aged 15-19 have had a child[58]. While a majority of women do work in Mwanza, only 34% of working women reported having control over how their income was used[58]. Women in Mwanza report a lower rate of female circumcision than the general Tanzanian population (1% vs 15%), but higher rates of physical violence (44% vs 39%) and sexual violence (25% vs 17%)[58]. In general, 100% of Tanzanian children enrol in and complete primary school, but less than half progress to secondary school[58].

- Technological: Tanzania is a technologically advancing country in which 75% of the population owns a mobile phone, a number likely higher in urban areas and among younger populations[42]. Further, even those who do not personally own mobile phones have access to one through the community[59]. Fifty-five percent of Mwanza's slum-dwelling population report using electricity at home for lighting[60].
- Environmental: Mwanza is a coastal city on the banks of Lake Victoria, a fresh water lake and the city's primary water source. Mwazna has clear seasonality, rainy periods between October and December and March and May. Mwanza has a generally flat topography, with much of its land being suitable for agricultural activities[61].

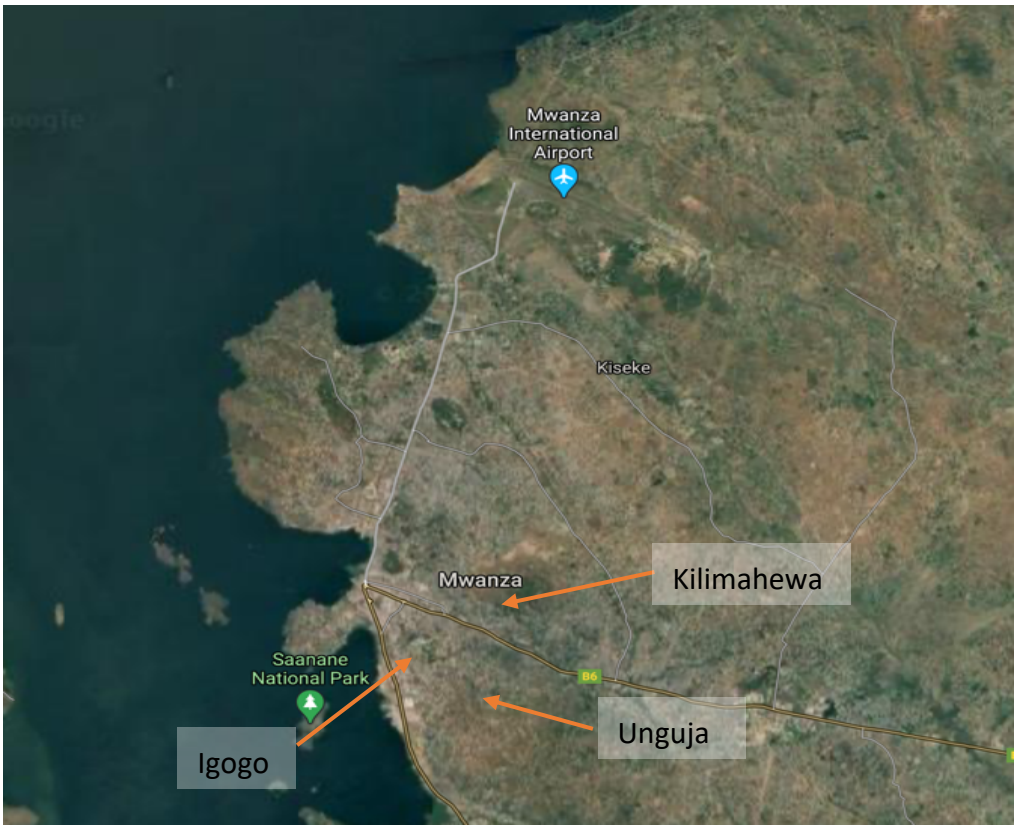


Figure 4: Locations of urban informal settlements selected as my study sites in Mwanza. From Google Maps, 2020

2. Estimating Variability Introduced by Current Diarrhoea Surveillance Methods: The Impact of Diarrhoea Surveillance Methods for Under-Fives in LMICs on Estimated Epidemiology: A Systematic Review and Meta-Analysis of Methodological and Primary Empirical Studies

2.1.Abstract

Background

Diarrhoea is the second highest cause of under-five death globally. Estimation of under-five diarrhoea rates is necessary to evaluate interventions. However, differences in findings across studies might reflect differences in the methods used rather than true diarrhoea rates.

Methods

I systematically reviewed all studies published between 2000 and 2018 that estimated under-five diarrhoea rates in low and middle income countries and extracted data on diarrhoea rates, surveillance methods, and reactivity. I summarised data from studies that performed direct comparisons of methods, and indirectly compared studies which utilised only one method using meta-regression to determine the association between methods and estimated diarrhoea rates.

Findings

Two-hundred-and-seventy-seven studies met my inclusion criteria: four direct comparisons and 273 studies utilising only one estimation method. Meta-regression across all studies showed that diarrhoea rates were sensitive to method used. I estimated that passive surveillance methods were associated with a 96% lower estimated rate than active surveillance (IRR=0·04,95%CI[0·02,0·09]). Among active surveillance studies, a doubling of recall period was associated with a 47% lower rate (IRR=0·53[0·46,0·60]), while a decreased questioning frequency was associated with a higher estimated rate: at the extreme, one time questioning yielded a nearly 4X higher rate than daily questioning (IRR=3·84[2·48,5·96]).

Interpretation

Estimated diarrhoea rates are sensitive to their estimation methods. There is a need for a standardisation of diarrhoea estimation methods, and for the use of other outcomes in the estimation of population level gastrointestinal health.

2.2.Introduction

As I have discussed in Chapter One, effective surveillance of diarrhoea in children under-five is required to track outbreaks, allocate public health resources, and evaluate WASH interventions[1]. However, the choice of method used estimate diarrhoea rates has been hypothesized to introduce undue variation in rates[18, 20]. If true, the results of observational and interventional studies may differ by method, thereby obscuring the effects of the explanatory variable(s) of interest. Since diarrhoea is one of the biggest causes of death in children, this is a methodological point of considerable practical importance. The two most common methods of diarrhoea surveillance used to estimate rates, passive and active surveillance, take different approaches. Passive surveillance relies on data collected from health facilities and therefore excludes all children with diarrhoea who do not attend a facility. Passive surveillance estimates are therefore skewed towards more severe disease and away from marginalised groups such as slum-dwellers, refugees, and migrants, who are less likely to visit health facilities and who are more likely to visit informal facilities than non-marginalised groups[23, 27, 28 29-31]. Passive surveillance is a useful, inexpensive tool to detect new outbreaks of severe diseases, such as cholera, but is arguably less useful as an epidemiological tool for the estimation of population-level diarrhoea rates or in trials of WASH interventions.

Active surveillance, based on door to door surveys, provides a more complete report of diarrhoea rates than passive surveillance, but may also be subject to measurement error and bias[23]. Carers may forget events that happen in the past, particularly during longer lengths of recall[33, 62]. They may also have a poor understanding of what diarrhoea is[37]. UNICEF and the Demographics and Health Surveys (DHS) program have recommended a method based on asking carers if their child has had three or more loose or watery stools in any 24-hour period within the previous 14 days[63]. However, this method is by no means universally applied. In addition to concerns over error and bias in estimates, concern has been

expressed that diarrhoea rates may be 'reactive,' in the sense that net of any true clinical difference, people report different diarrhoea rates for psychological reasons – either because they perceive themselves to be beneficiaries of an intervention, or because the surveyors provided subliminal clues[36, 64]. As such, reactivity creates a particular concern for evaluations of WASH interventions.

In order to examine the association between the method used in diarrhoea surveillance and estimated diarrhoea rates I conducted a systematic review of all studies published between 2000 and 2018 that report under-five diarrhoea rates in LMICs. I examined studies that performed direct 'head to head' comparisons of different methods in order to estimate differences in diarrhoea rates by method. However, I found only four such studies. I therefore obtained studies that used only one estimation method so that I could compare the estimated diarrhoea rates of each method across studies indirectly by means of meta-analytical methods[65, 66].

The aims of the systematic review and meta-analysis were to determine: 1) the frequency of use of the different diarrhoea rate estimation methods; 2) the association between passive and active surveillance methods and estimated diarrhoea rates; 3) the association between a. recall periods, b. questioning frequencies, and c. prospective (diary) vs retrospective recall on estimated diarrhoea rates among active surveillance studies; and 4) the extent of reactivity in diarrhoea rate estimation.

2.3.Methods

The literature retrieval for this study followed PRISMA guidelines, and was registered on PROSPERO (CRD42018119062)[65, 67].

Search Strategy

I conducted a systematic review of studies published between 2000 and 2018 that made quantitative estimates of diarrhoea rates among children under the age of five in LMICs (as defined by the Organisation for Economic Co-operation and Development)[68]. The search strategy aimed to capture any study that estimated diarrhoea rates among under-fives, including both studies that performed direct 'head to head' comparisons of methods, and studies estimating diarrhoea rates using only one method that I can compare indirectly. I excluded studies that were not designed to estimate population level diarrhoea rates, such as studies which measured hospital acquired infection; clinical trials in which diarrhoea was an adverse drug event; and case-control studies in which diarrhoea was the case. Studies were restricted to English or French (Table 1).

Table 1 Inclusion and Exclusion Criteria for the Systematic Review

Inclusion Criteria	**Exclusion Criteria**
Reports diarrhoea epidemiology among under-fives (i.e. incidence or prevalence)	Nosocomial setting
Takes place in LMICs	Clinical trials in which diarrhoea was an adverse drug event
Published between 2000 and 2018	Case-control studies in which diarrhoea is the case
English or French language	

I searched the MEDLINE, Embase, and PubMed databases for studies matching the inclusion and exclusion criteria. The search string included key words relating to diarrhoea and population-level disease measurement and estimation. The string

then restricted the results to human studies and studies in LMICs. The full search string can be found in Appendix I.

SW and RR independently screened each title and abstract, and any disagreements were resolved through discussion with RL. Full texts were screened in a similar manner. Where full texts were not available, I requested the article from the University of Warwick's Article Reach Service. I excluded unavailable studies and duplicate studies. In the event that multiple studies used the same data source, I selected a random study for inclusion.

Data Extraction

RR extracted the data, with SW duplicating data extraction for random 15% of the included studies. There was full agreement between RR and SW. As a single reported study may have included several independent estimates of diarrhoea rates, I treated each report as separate 'estimates' within one study. This would apply to all the direct studies but also arose in indirect studies when conducted in more than one site, included multiple rounds of data collection and/or was a trial with multiple arms. For example, a two-armed trial which estimated diarrhoea rates at base-line and end-line yielded four estimates.

Table 2 reports the data I extracted. Data extracted included participant demographics, study design, diarrhoea rates, surveillance methods used to estimate rates, and whether or not the study was a direct comparison of methods. I defined direct comparison studies as those which included at least two separate arms, each with a different method of diarrhoea estimation (including altering recall period or questioning frequency) that compare estimated diarrhoea rates between each arm.

Table 2 Data extracted from each study in the Systematic Review

Variable	**Units (if applicable)**	**Categorisation (if categorised)**
Authors	NA	NA
Study Title	NA	NA
Year of Publication	Years	NA
Year of Data Collection	Years	NA
Mean Participant Age	Years (with Standard Error)	NA
Participant Sex Breakdown	% Female	NA
Geography	Urban, Rural, or Mixed	NA
Country	NA	Regions as defined by the United Nations Development Program (UNDP)[69]
Intervention (for RCTs)	NA	WASH, Health, or Nutrition
Effectiveness of Intervention (for RCTs)	Incidence Risk Ratio	NA
Diarrhoea Rate (in incidence)	Episodes per child-year (converted if necessary)	NA
Surveillance Method	NA	Passive, Retrospective Active, Prospective (Diary) Active
Recall Period	Days	NA
Questioning Frequency	Days	Daily, Weekly, Monthly, or Annually or longer
Questioning Type	NA	Pictorial, Verbal, or other
Sample Size (observed cases for passive surveillance studies)	NA	NA
Study Design	NA	Observational (Primary), Observational (Secondary), RCT (broken down by arm)
Direct Comparison of Methods	NA	Yes/No

Data Analysis

I indirectly compared the diarrhoea rates from the included estimates, including those from studies that performed direct comparisons of estimation methods, and studies using only a single method of estimation. I summarised the key variables in each individual estimate, including estimated diarrhoea rates, surveillance method, year, and region (as defined by UNDP). I estimated a hierarchical meta-regression model for log diarrhoea rate (incidence in episodes per child-year), adjusting for region and time (in years) and interactions between time and region. I also adjusted for study design, including a categorical variable with levels: 1) Observational studies which use primary data sources, 2) Observational studies which use secondary data sources, 3) RCT intervention arms before the intervention, or the control arm (if reported), and 4) RCT experimental arm after the intervention. The model was estimated in StataSE Version 15 using generalised least squares with random effects at the study level, to account for within-study correlation, and weighting by the study sample size[69, 70].

I estimated two separate meta-regression models. The first iteration included the estimates from both passive and active surveillance studies to estimate the effect of surveillance types (passive/active) on estimated diarrhoea rates, and as such included a dummy variable for surveillance type. I also estimated pooled temporal and regional trends from this model.

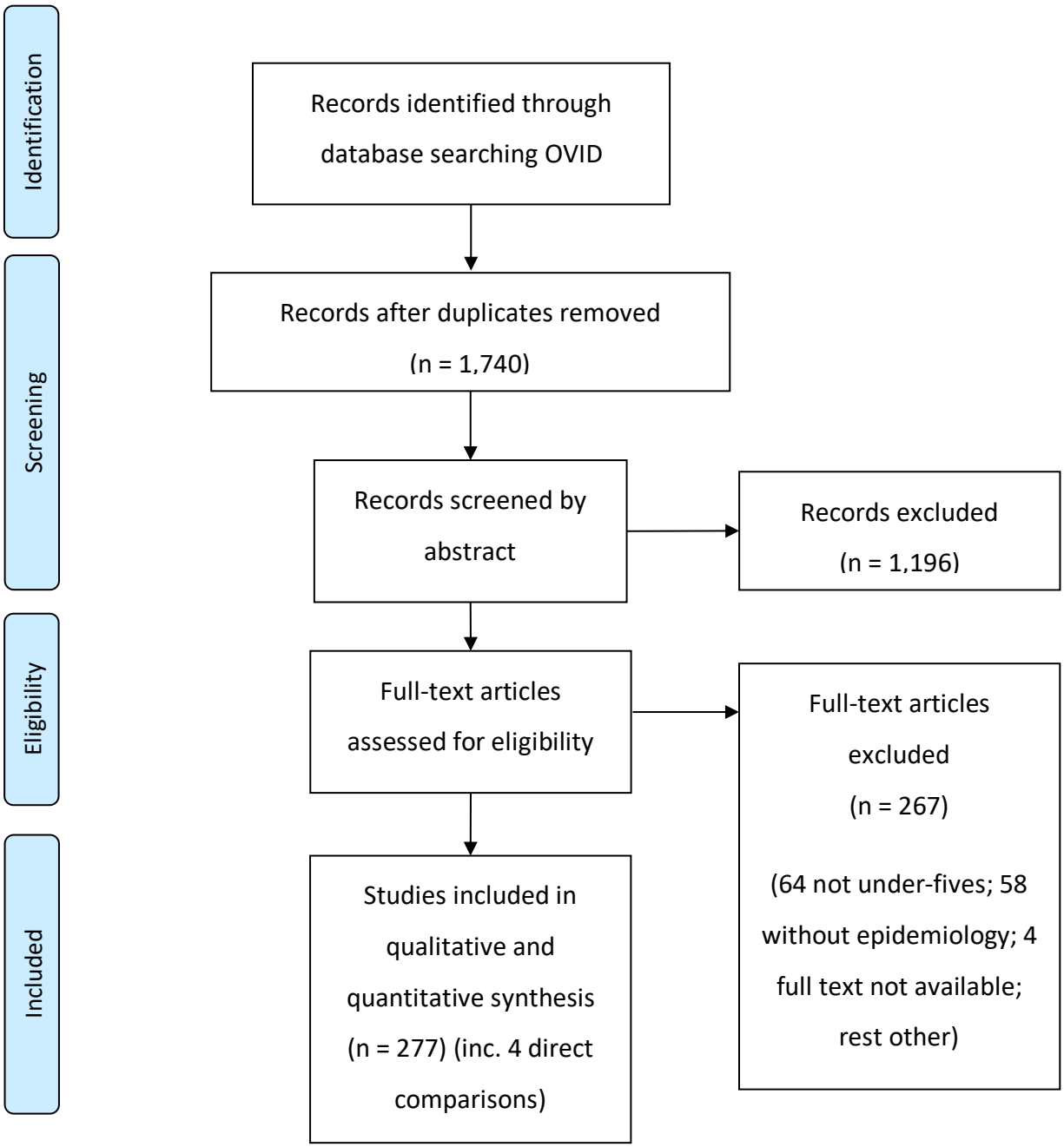
The second iteration included only estimates from active surveillance studies to examine the effects of variables exclusive to active surveillance studies on estimated diarrhoea rates. These included variables for recall period (as a continuous numeric term in days), and questioning frequency and recall type (both as categorical variables). Further, any other variations found between active surveillance studies, including reactivity and questioning type (e.g. verbal or pictorial) were included.

2.4.Results

Study Identification

I identified 1,807 studies in total, which was reduced to 1,740 after duplicates were removed. Abstract and title screening yielded 544 studies, with a further 267 excluded after full-text review. Common reasons for exclusion included not presenting data on under-fives (n=64), not including data on diarrhoea rates (n=58), and not being able to obtain the full text (n=4). Overall, 277 full text studies were included in the final review (Fig. 5) (studies are listed in Appendix II).

Figure 5 PRISMA Diagram of the Systematic Review



Study Characteristics

As stated, many studies included more than one estimate of diarrhoea rate, arising from estimates at different time points, being a trial with two or more arms, or data collection in more than one location. Appendix III presents information on the number of estimates per study. I identified 612 separate estimates of population-level diarrhoea rates, and these constitute my denominator. In total there were 591 (97%) active surveillance estimates and 21 (3%) passive surveillance estimates. Of the 591 active surveillance estimates, 578 (97%) were retrospective while 13 (3%) were prospective (a diary kept by the carer). Three-hundred and four (49%) of the estimates used a 14-day recall period, as recommended by UNICEF and the demographic and health survey (DHS) program. Of the 188 estimates which came from randomised control trials (RCTs) 53 (28%) used a 14-day recall period and 95 (51%) used a 7-day recall period. Further, of the estimates from RCTs, 21 (12%) questioned daily, 46 (24%) questioned weekly, and 49 (26%) questioned two-weekly.

By region (as defined by the UNDP), Sub-Saharan Africa was the setting for the largest number of estimates (197; 32%), followed by South America (170; 28%), East and South-East Asia (118; 19%), Central and Southern Asia (112; 18%), and finally North Africa and the Middle East (15; 2%). Rural (218; 41%) and mixed geography (188; 35%) areas provided more estimates than urban areas (124; 23%).

None of the included active surveillance estimates used non-verbal methods of diarrhoea estimation (e.g. showing carers pictures of stool), and no studies made mention of a 'gold standard' of diarrhoea estimation.

Four of the included studies performed direct head-to-head comparisons of diarrhoea estimation methods: three examining the effect of differing recall periods on diarrhoea rates, and one examining the effect of questioning frequency on

estimated diarrhoea rates. No studies were identified that analysed reactivity in diarrhoea rate estimation.

Table 3 Summary table of the 604 estimates included in the systematic review, with average estimated diarrhoea rates for key variables

		Passive n(% of total)	**Prospective (Diary) Active** n(% of total)	**Retrospective Active** n(% of total)	**Total**	**Average reported diarrhoea rate (Episodes per Child-Year)**
	TOTAL	21	13	578	612	6·93
Data Collection Year	2016-2017	0 (0%)	0 (0%)	6 (100%)	6	19.67
	2011-2015	10 (6·5%)	2 (1·3%)	141 (92.2%)	153	7.01
	2006-2010	5 (2·8%)	2 (1·1%)	177 (96·2%)	184	6.62
	2001-2005	4 (3·5%)	2 (1·8%)	111 (94·9%)	117	7.82
	1996-2000	2 (2·6%)	1 (1·3%)	75 (96·2%)	78	5.96
	1991-1995	0 (0%)	0 (0%)	38 (100%)	38	5.99
	1986-1990	0 (0%)	3 (42·9%)	4 (57·2%)	7	3.53
	Missing Data	0 (0%)	3 (10·4%)	26 (89·7%)	29	6.92

Region	Sub-Saharan Africa	3 (1·5%)	3 (1·5%)	191 (97.0%)	197	6.45
	North Africa and the Middle East	0 (0%)	1 (6·7%)	14 (93·4%)	15	4.97
	Central and Southern Asia	2 (1·8%)	2 (1·8%)	108 (96·5%)	112	10.06
	Eastern and South-East Asia	11 (9·4%)	4 (3·4%)	103 (87·3%)	118	6.22
	The Americas	5 (3%)	3 (1·8%)	162 (95·3%)	170	6.10
Study Design	Observational (Primary)	16 (5·6%)	11 (3·9%)	257 (90·5%)	284	8·16
	Observational (Secondary)	5 (3·6%)	0 (0%)	135 (96·4%)	140	3·91
	RCT (Baseline or Control)	0 (0%)	1 (1·1%)	93 (98·9%)	94	8·37
	RCT (Experimental Post-Intervention)	0 (0%)	1 (1·1%)	93 (98·9%)	94	6.31
Geography	Mixed	10 (5·4%)	2 (1·1%)	176 (93·7%)	188	4.13

	Rural	4 (1.8%)	4 (1.8%)	210 (96·3%)	218	10.54
	Urban	4 (3·3%)	7 (5·7%)	113 (91·2%)	124	4·79
	Missing Data	3 (3·7%)	0 (0%)	79 (96·4%)	82	6.98
Questioning Frequency	Daily	NA	0 (0%)	55 (100%)	55	4.68
	Weekly	NA	9 (9·7%)	84 (90·4%)	93	4.35
	Monthly	NA	1 (1·3%)	77 (98·8%)	78	13.47
	Annual+	NA	0 (0%)	30 (100%)	30	5·43
	One Off	NA	3 (1%)	332 (99·1%)	335	7.05
Recall Period	1-3 days	NA	1 (1·7%)	59 (98·4%)	60	25.62
	4-14 days	NA	11 (2·4%)	468 (97·7%)	479	5.33
	15-30 days	NA	1 (2·8%)	35 (97·3%)	36	3.68
	31-90 days	NA	0 (0%)	13 (100%)	13	1.00
	91+ days	NA	0 (0%)	3 (12·5%)	24	1.83

Differences between Active and Passive Surveillance

No studies performed direct 'head to head' comparisons of active and passive surveillance. After model-based adjustment to perform an indirect comparison of passive and active surveillance, passive surveillance was associated with a 96% lower estimated diarrhoea rates than active surveillance (Incidence Risk Ratio (IRR)=0·04[95%CI:0·02,0·09]) (Table 4).

The Impacts of Factors Within Active Surveillance Studies on Estimated Diarrhoea Rates

Questioning Frequency

One study directly examined the effect of differing questioning frequencies on estimated diarrhoea rates; Zwane et al. *(*2011) estimated that biweekly surveys estimated a 7-15% lower diarrhoea rate than six-monthly surveys when using the same recall period[71].

My indirect comparison of active surveillance estimates produced comparable results; after model-based adjustment, I found that that less frequent questioning was associated with an increase in estimated diarrhoea rates. For example, one-time questioning was associated with a rate close to four times higher than daily questioning (IRR=3·84[2·48,5·96]) (Table 4). This is not evident graphically in unadjusted crude data – however, a large amount of variance as questioning frequency increases can still be seen (Fig. 6)

Recall Period

Three studies directly examined the effect of differing recall periods on estimated diarrhoea rates, although the recall periods examined were different. Melo et al. (2007) found that diarrhoea rates were cut by a third when carers recall over 4 weeks compared to 24 hours[72]. Feikin et al. (2010) similarly estimated that diarrhoea rates were cut by a fifth when carers recall over 11-13 days compared to 1-2 days[33]. Lee et al. (2010), estimated that estimated diarrhoea rates were similar

for carers who recalled over a 72-hour period compared to a 24-hour period[73], but this is a much shorter range than that investigated in the other two studies.

Based on an indirect comparison of the included active surveillance estimates, I estimated that recall periods and estimated diarrhoea rates were inversely associated. After model-based adjustment, I found that a doubling of recall period was associated with a 47% reduction in diarrhoea rate (IRR=0·53[0·46,0·60]; Table 4). This is also evident graphically in the crude data (Fig. 7).

Prospective Vs Retrospective

No studies directly compared prospective (diary) recall designs against retrospective. I estimated through indirect comparison that retrospective recall estimates were associated with a lower rate than prospective (diary) estimates, but the effect size was relatively uncertain (IRR=0·66[0·41,1·07]; Table 4).

Table 4 A random effects model showing the association between study characteristics and reported diarrhoea rates for all studies and only active surveillance studies in the Systematic Review

		Active Vs Passive		**Within Active**	
		IRR	**95%CI**	**IRR**	**95%CI**
Surveillance Type	Reference: Active Surveillance	Ref.	Ref.	Ref.	Ref.
	Passive Surveillance	0·04	(0·02,0·09)	NA	NA
Region	Reference: Sub-Saharan Africa	Ref.	Ref.	Ref.	Ref.
	North Africa/Middle East	0·08	(0·06,0·10)	0·09	(0·06,0·13)
	Central and South Asia	1·96	(0·04,96·96)	1·92	(0·05,90·44)
	Eastern and South-East Asia	0·61	(0·26,1·49)	0·63	(0·26,1·53)
	South America and Caribbean	0·81	(0·62,1·06)	1	(0·73,1·39)
Year	Year	0·97	(0·96,0·98)	0·97	(0·97,0·98)
Region and Year	Reference: Sub-Saharan Africa	Ref.	Ref.	Ref.	Ref.
	North Africa/Middle East	1·14	(1·12,1·15)	1·13	(1·11,1·15)
	Central and South Asia	0·98	(0·84,1·15)	0·99	(0·85,1·15)
	Eastern and South-East Asia	1·03	(0·99,1·06)	1·03	(0·99,1·06)
	The Americas	1·02	(1·01,1·04)	1·01	(1,1·03)
Study Design	Reference: Observational Primary	Ref.	Ref.	Ref.	Ref.
	Observational Secondary	1·01	(0·83,1·22)	1	(0·83,1·22)
	RCT (Baseline or Control)	1·16	(0·84,1·6)	1·55	(1·11,2·15)
	RCT (Experimental Post-Intervention)	1·01	(0·75,1·36)	1·37	(0·98,1·91)
	Log of Measurement Period	NA	NA	0·53	(0·46,0·60)
Questioning Frequency	Reference: Daily	NA	NA	NA	NA
	Weekly	NA	NA	1	(0·62,1·62)
	Monthly	NA	NA	2·16	(1·14,4·08)
	Annually or Longer	NA	NA	4·66	(2·6,8·35)
	One Off	NA	NA	3·84	(2·48,5·96)
Method	Reference: Diary	NA	NA	Ref.	Ref.
	Self Report	NA	NA	0·66	(0·41,1·07)
	Constant	7·22	(5·7,9·15)	17·49	(9·69,31·58)
Random Effects Parameters	Variance (Constant)	1·01	(0·80,1·28)	0·68	(0·47,1·00)
	Variance (Residual)	0·07	(0·05,0·09)	0·063	(0·05,0·08)

Figure 6 An unadjusted box plot for estimated diarrhoea rate against questioning frequency for active surveillance studies

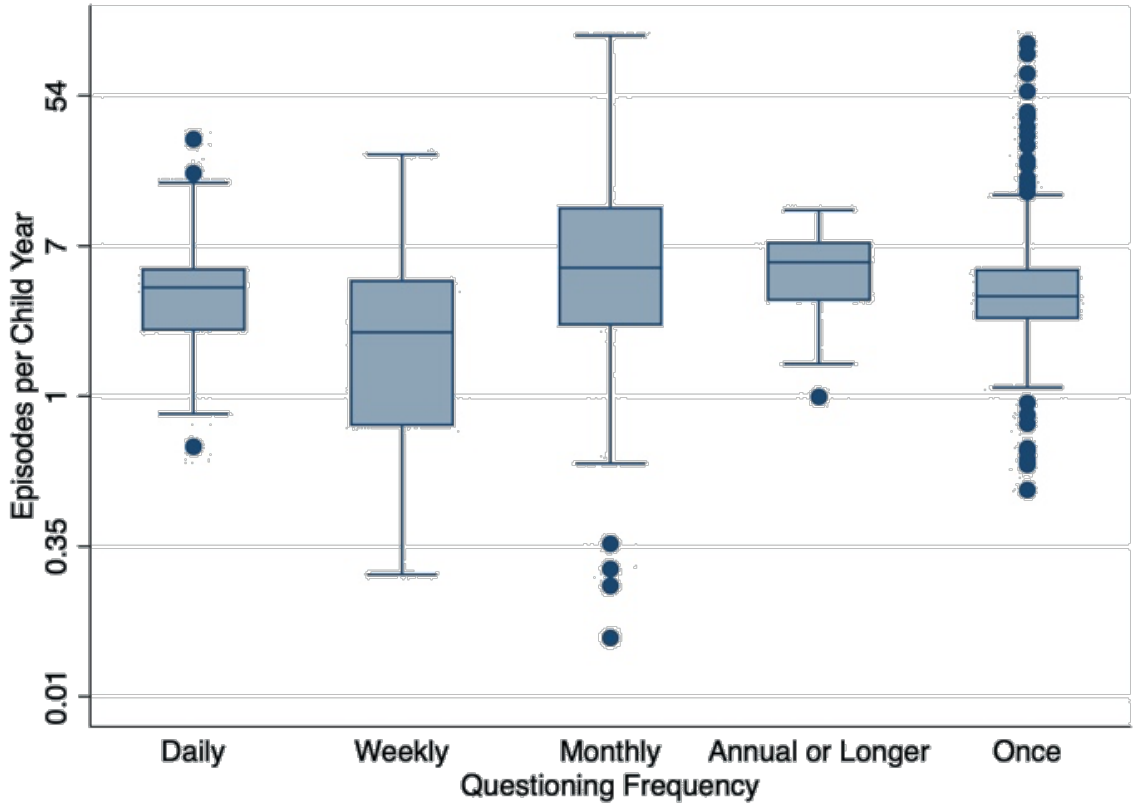
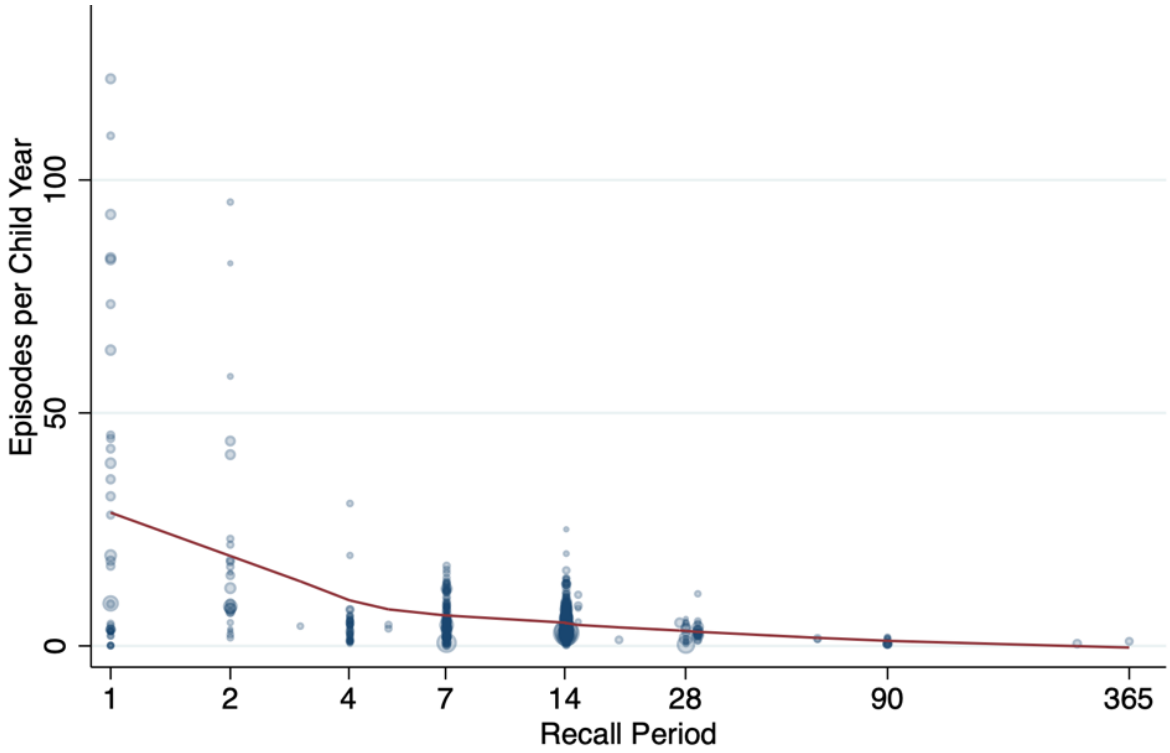


Figure 7 An unadjusted scatter plot and trendline for estimated diarrhoea rate against recall period for active surveillance studies



2.5.Discussion

Main Findings

I provide evidence that estimates of under-five diarrhoea rates are sensitive to the methods used. This includes variance introduced by the choice of passive or active surveillance, as well as factors specific to active surveillance.

Passive surveillance methods were associated with 96% lower diarrhoea rates than active surveillance methods. The most probable explanation is that carers do not seek health care for the majority of cases where, if asked, they would report diarrhoea. While not shown in my results, several studies on access to healthcare among infants in LMICs show that the propensity of carers to seek care for their under-fives with diarrhoea is influenced by diarrhoea severity, socioeconomic or legal status, and other demographic characteristics[27-31].

Regarding active surveillance methods, I found that different questioning frequencies influence estimated diarrhoea rates. There was a trend to lower estimated diarrhoea rates given higher questioning frequencies: one-off questioning was associated with a nearly four times higher estimated diarrhoea rate than daily questioning. I also found that differing recall periods were associated with a change in estimated diarrhoea rates: a doubling of recall period was associated with a halving of estimated diarrhoea rates.

Factors that Result in Subjective Estimates during Active Surveillance

As the distinction between the recall of a diarrheal or non-diarrheal stool by carers is largely subjective, several cognitive factors can affect estimates. These include respondent fatigue (becoming tired of answering questions), recall bias (forgetting events that have occurred in the past), perception bias (not understanding the question being asked), and reactivity (answering differently due to experiencing an intervention).

Respondent Fatigue

Declining diarrhoea rates with increasing questioning frequencies (but the same recall period) suggest respondent fatigue – participants may be inclined to pay attention to their bowel movements at first but lose motivation with further rounds of questioning.

Recall Bias

Recall bias is the effect of forgetting: participants are more likely to recall recent than older events. I would expect a lower reported number of diarrhoea episodes with longer recall periods and this was borne out by my analysis including two of the three head-to-head comparisons – the exception examined a much smaller gap between questions than the other two[33, 74]. This finding was corroborated by my more recent head-to-head comparison in the next chapter (not included in the review because it was published after 2018), where daily recall was associated with a 30 percentage point higher estimated diarrhoea rate than fortnightly recall during a text message survey of under-five diarrhoea in urban Tanzania[75]. While not examined in my review, it has also been reported that the effect of recall bias is more apparent for moderate diarrhoea compared to severe diarrhoea. Zafar et al. (2010), for example, found that moderate diarrhoea is reported at half the rate of severe diarrhoea during longer recall periods[62].

Other factors affecting diarrhoea estimation

Other factors outside the scope of this review can further influence reported diarrhoea rates. For example, poor caregiver perception of diarrhoea (understanding what is or is not diarrhoea) can result in error in diarrhoea estimation. Voskuijl et al. (2017) determined that carers of children under five were only able to identify 56%-75% of loose or watery stools, and 80% healthy stools[37].

Another relevant phenomenon is “reactivity” whereby participants adjust their answers to a survey according to how they believe they ought to respond, regardless of any true underlying difference. I did not identify any studies of reactivity in this review but it has been discussed as a potential explanatory factor in previous trials. Luby et al. (2018) stated that ‘people who received the intervention might have been grateful and, out of courtesy, reported less diarrhoea’[36]. Wood et al. (2008) further found evidence for reactivity in clinical trials for various diseases, reporting that inadequate concealment of interventions is associated with improved treatment performance in trials, particularly for subjective outcomes[76].

Implications

The magnitude of the variation in diarrhoea rates, even among active methods of surveillance, suggests the need for standardization of diarrhoea estimation methods to facilitate comparisons between studies. Despite the fourteen day UNICEF and DHS standard, there does seem to be a trend towards using a 7-day recall period. It was the most frequently used recall period (51%) among RCTs in my review. Three of the three recent large integrated WASH trials (the SHINE trials in Bangladesh and Kenya, and the WASH Benefits trial in Zimbabwe) also used a seven day retrospective recall to measure diarrhoea, in contravention of the UNICEF and DHS guidelines[77-80]. However, the above three trials differed among themselves with respect to question frequency: the SHINE trials questioned carers annually, while the WASH Benefits trial questioned mothers every ‘two to six’ months[78-80]

It could be argued that lack of standardisation simply introduces error in estimates that can be counteracted by increasing sample size. However, this is only likely to be true if it is assumed that the different methods affect only the propensity of someone with a true case of diarrhoea (or indeed enteric infection) to report a case of diarrhoea (the “sensitivity” of the method)[81]. If, however, there is a loss of

"specificity" – the propensity of someone who did not have diarrhoea (or enteric infection) to report a case of diarrhoea – then intervention effects will also be biased across studies using different methods. It is also likely that any estimation errors will bias results towards the null, rather than towards reports of intervention effectiveness[81]. It is therefore possible that the choice of methodology is at least partially responsible for the widely varying, and often disappointing, results of evaluations on WASH interventions[19, 36, 78, 80].

While a widely accepted standard would facilitate comparisons across different observational and experimental studies, this raises the question of what the optimal standard may be that would also produce reliable reports of diarrhoea rates. There is no "gold standard" method for estimating of diarrhoea rates. In part, this is because of the difficulty in defining the underlying construct and providing a culturally and linguistically consistent definition of a case or episode of 'diarrhoea'. Direct observation by an expert might constitute a gold standard against which other methods could be compared, as has been described above in the study by Voskuiljl et al. (2017)[37]. However, judgements among experts may not be universal. Moreover, the collection of every stool and use of experts to classify them quickly becomes impractical at larger scales. Nonetheless, we examine this in Chapter Four.

I propose two policies to mitigate the problem. First, an agreed consensus method for the estimation of diarrhoea rates in surveys. This consensus method may be similar to what is already done, or a new method, including the use of new technologies. I explore one such technology, text message surveys, in the next chapter. Second, triangulation of diarrhoea rates with other observations that reflect on gastrointestinal health when interventions are evaluated. Many WASH evaluations already include anthropometric measurements as outcomes alongside diarrhoea. Further, direct measurement of environmental contamination and pathogen levels in stool samples should complement diarrhoea rates in clinical

studies. This would also allow for determination of how much diarrhoea is attributable to infection (and which can be reduced by WASH interventions), rather than non-infective reasons (which would likely not be impacted by WASH interventions). Investigation of the link between interventions, environmental contamination, and the profiles of pathogen carriage in childhood stools is an important topic for scientific research. I also explore this in Chapter Four.

3. Methods to Decrease Variation in Estimates of Diarrhoea Rates: Effectiveness of Text Message Surveys for Diarrhoea Estimation: A Factorial Cross-Over Randomised Controlled Trial

This chapter has been published in its entirety: Rego, R., Watson, S., Ishengoma, P. et al. *Effectiveness of SMS messaging for diarrhoea measurement: a factorial cross-over randomised controlled trial. BMC Med Res Methodol 20, 174 (2020). https://doi.org/10.1186/s12874-020-01062-3*

3.1.Abstract

Background

Text messaging systems can be used to estimate disease prevalence. Using a text messaging system, I evaluated the effects of question load, question frequency, and financial incentive on response rates and estimated infant diarrhoea rates in an infant diarrhoea survey.

Methods

I performed a factorial cross-over randomised controlled trial of an SMS surveying system for infant diarrhoea surveillance with treatments: financial incentive (yes/no), question load (1-question/3-question), and questioning frequency (daily/fortnightly). Participants progressed through all treatment combinations over eight two-week rounds. Data were analysed using multivariable logistic regressions to determine the impacts of the treatments on the response rates and estimated diarrhoea rates. Attitudes were explored through qualitative interviews.

Results

For the 141 participants, the mean response rate was 47%. In terms of percentage point differences (ppd), daily questioning was associated with a lower response rate than fortnightly (-1·2[95%CI:-4·9,2·5]); high (3-question) question loads were associated with a lower response rate than low (1-question) question loads (-7·0[-10·8,-3·1]); and financial incentivisation was associated with a higher response rate than no financial incentivisation (6·4[2·6,10·2]).

The mean two-week estimated diarrhoea rate was 36·4%. Daily questioning was associated with a higher estimated diarrhoea rate than fortnightly (29·9[22·8,36·9]); with little evidence for impact by incentivisation or question load.

Implications

Close to half of all participants responded to the SMS survey. Daily questioning evoked a statistically higher rate of estimated diarrhoea, while financial incentivisation and low (1-question) question loads evoked higher response rates than no incentive and high (3-question) question loads respectively.

3.2.Introduction

Data from infectious disease surveillance systems are used in numerous ways, including the evaluation of public health interventions, providing early warning of outbreaks, and allowing for the proper allocation of resources. For children in low and middle income countries (LMICs), diarrhoeal disease is a key surveillance target, as diarrhoeal disease is globally the second highest cause of under-five mortality, as well as a large contributor to stunting and perhaps cognitive delay[1, 3]. There are two basic methods of disease surveillance: passive surveillance, such as the WHO's EWARS based on reports of targeted diseases from health facilities; and active surveillance, where households are visited randomly to ascertain diarrhoea rates. Both methods have strengths and weaknesses. In my systematic review, I found that passive surveillance is inexpensive but underestimates diarrhoea rates (as not all affected patients visit reporting health facilities) (Chapter Two). On the other hand, active surveillance, which is based on door to door questioning, may detect a higher proportion of cases but is costly and time-consuming (Chapter Two). Factors within active surveillance methods may also introduce variation in estimated rates.

Ninety percent of people in LMICs have access to basic mobile phones which can be used for health promotion and disease surveillance[41, 42]. In Southern Africa, for example, mobile phones have been used effectively for over a decade to ensure adherence to antiretroviral medication for HIV[82]. Mobile phones have also been used in Ghana during a demographic and health survey, and during the 2014-2015 West African Ebola outbreak[44, 45]. During the 2014-2015 Ebola outbreak in Liberia, mobile phone surveys were deployed to find Ebola cases, estimate Ebola mortality rates, and evaluate care-seeking behaviours of Ebola patients[44]. The Liberian study, utilising a random list of phone numbers, conducted both phone call and text message surveys. The study received a response rate of 15% to text messages and 13% to phone calls[44]. The results also showed a significant drop in response rates

for both messages and phone calls between the first and second rounds of data collection – from 22% to 11% for text messages and 18% to 10% for phone calls[44]. L'Engle et al. (2018) used mobile phone voice message surveys to measure demographics and health behaviour in a Ghanaian non-emergency setting, obtaining a response rate of 31%[45]. The study also found that younger, urban, highly educated, and male respondents were more likely to respond to the mobile phone survey than face to face surveys – possibly resulting in selection bias[45].

It is clear that mobile phones are capable of estimating disease rates. However, response rates in past studies were not very high and were influenced by numerous factors. Some factors may relate to the design of the system. I therefore decided to investigate the effect of certain factors of design on response rates. Given the importance of childhood diarrhoea, I selected this as the disease of interest for my study. The estimation of rates of childhood diarrhoea is also subject to a considerable amount of error given factors such as recall period and questioning frequency, with a majority of studies choosing either 24-hour recall or 14-day recall (Chapter Two). I therefore decided to measure the effects financial incentivisation (yes/no), recall period/questioning frequency (24-hour vs 14-day), and question load (1-question/3-question) on response rates and estimated diarrhoea rates in an SMS survey.

3.3.Methods

This trial is reported in line with the Consolidated Standards of Reporting Trials (CONSORT) Statement[83]. The CONSORT Checklist can be found in Appendix IV.

Study Design

I conducted a factorial, multiple crossover randomised control trial (RCT) of three SMS messaging data collection formats: daily vs fortnightly messaging, incentivisation of 1000TZS (~0·40USD) per response vs no incentivisation, and a high question load (1-question) vs a low question load (3-question) survey instrument. This resulted in eight possible combinations of formats. The participants progressed through all eight treatment combinations in a random order over eight two-week rounds, between April and September 2019.

The study took place in three informal settlements in Mwanza, Tanzania, where participants were recruited with the assistance of local community leaders. Community leaders assembled adults who cared for at least one child between 6 and 60 months, had access to their own mobile phone, and expressed interest in participating. Two or three meetings were held in each of the three communities where potential participants were invited to attend at a time of their convenience. The meetings provided potential participants with an opportunity to learn more about the study, ask questions, and discuss the project with the study team and their peers. Those who wished to participate provided written consent to field workers and were enrolled in the study. Consenting participants then completed a short demographic questionnaire.

Randomisation

Eight study arms were formed such that at any time point one arm would be receiving one of the eight treatment combinations, with no arm receiving the same treatment combination at any time point. Treatment sequences were randomly

generated by SW for each study arm with the restriction that no arm would receive the same incentive for more than two consecutive rounds or the same recall period for more than one round (Table 5). Further, each arm was sequenced to receive each of the eight treatment combinations over the study. RR then randomised participants at a 1:1:1:1:1:1:1:1 ratio into each arm using Microsoft Excel's RAND function. Participants were blinded to their arm allocations and sequence.

Table 5 Treatment combinations for each arm (A-H) during each study round with treatments

		Round 1			Round 2			Round 3			Round 4			Round 5			Round 6			Round 7			Round 8		
		F	I	Q	F	I	Q	F	I	Q	F	I	Q	F	I	Q	F	I	Q	F	I	Q	F	I	Q
Arm	A	F	N	H	D	I	L	F	I	H	D	N	L	F	N	L	D	I	H	F	I	L	D	N	H
	B	D	I	L	F	N	L	D	N	H	F	I	H	D	I	H	F	N	H	D	N	L	F	I	L
	C	D	N	L	F	I	L	D	I	L	F	N	H	D	N	H	F	I	H	D	I	H	F	N	L
	D	D	N	H	F	I	H	D	I	H	F	N	L	D	N	L	F	I	L	D	I	L	F	N	H
	E	F	I	H	D	N	L	F	N	H	D	I	H	F	I	L	D	N	H	F	N	L	D	I	L
	F	F	I	L	D	N	H	F	N	L	D	I	L	F	N	H	D	N	L	F	I	H	D	I	H
	G	F	N	L	D	I	H	F	I	L	D	N	H	F	I	H	D	I	L	F	N	H	D	N	L
	H	D	I	H	F	N	H	D	N	L	F	I	L	D	I	L	F	N	L	D	N	H	F	I	H

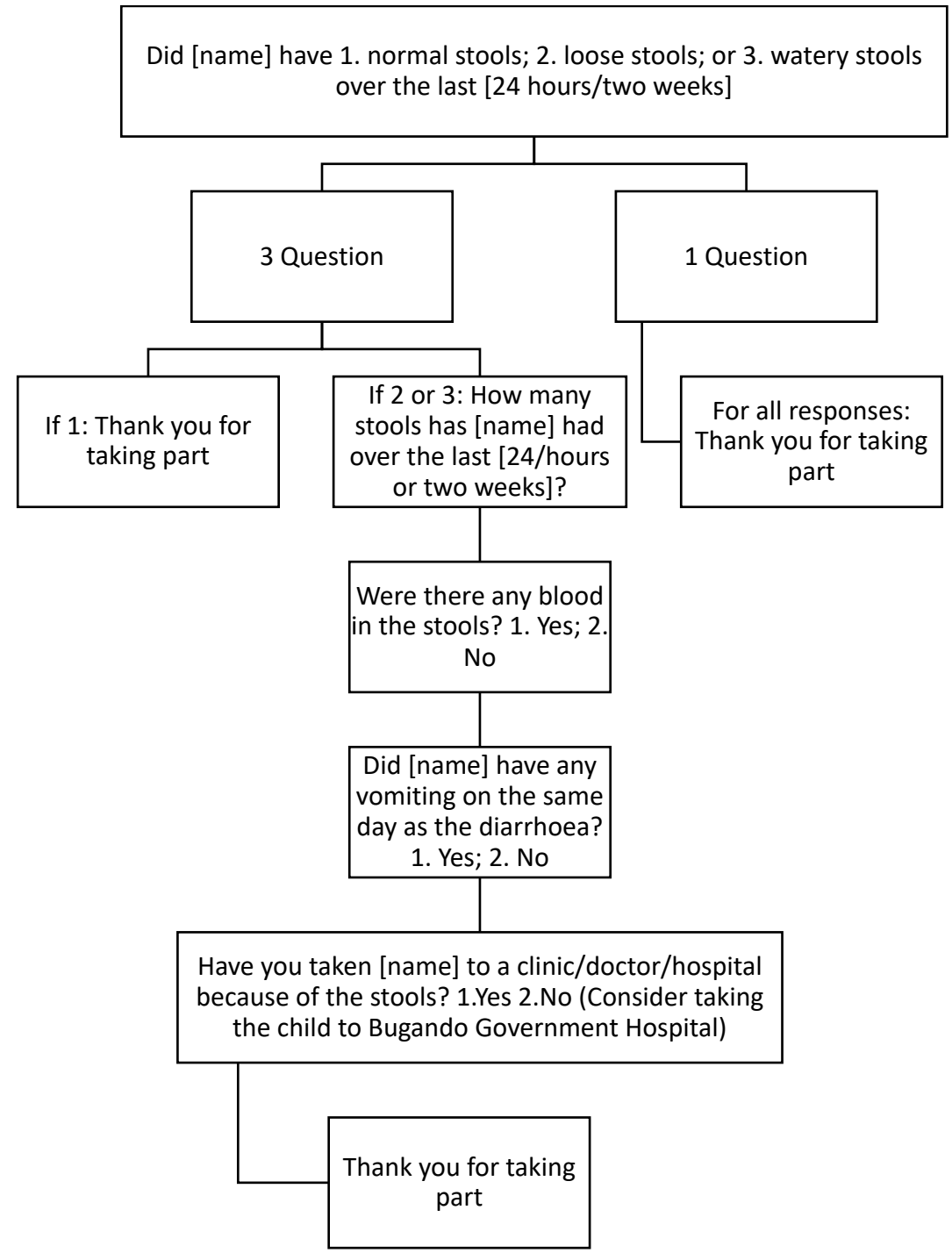
Table 5 Legend: 1) frequency (F), varying as daily (D) and fortnightly (F); 2) incentive (I), varying as incentive present (I) and no incentive (N); and 3) question load (Q), varying as high/3-question (H) and low/1-question (L)

Procedures

A text message, formatted according to the randomisation schedule, was sent via SMS message to participants between 10 AM and 11 AM local time on days due. On days due, all participants also received TZS 500 (~0·20USD) in airtime to cover the cost of responding to the survey. Participants receiving the incentive airtime payments were informed that this would be provided upon completion of all survey questions. If participants did not respond or complete the survey, they would receive two reminders – one after four hours, and the second after a further four hours. Responses were not accepted beyond 12 hours from the initial message. Participants choosing to participate were sent the applicable survey as per their assigned arm (Table 5). Participants receiving the 1-question survey were asked if their child had normal stool, loose stools, or watery stools over the past 24 hours or 14 days (dependent on frequency treatment) (Fig. 8). Participants receiving the 3-question survey were asked additional questions regarding blood in stool, vomiting, and health facility visits, if they reported loose or watery stool in the initial question. (Fig. 8).

I conducted nine qualitative interviews with carers taking part in the study to determine barriers and motivators to the SMS surveys. Nine were chosen as this was the number needed to reach theoretical saturation. Participants for qualitative interviews were chosen purposively to select those who had answered with different levels of frequency and to represent each study settlement. Sampling ensured that in each settlement, one person who never answered, one person who answered consistently, and one person who answered with varying consistency was interviewed. The interviews took place in KiSwahili and were conducted using a semi-structured guide exploring questions on daily life, attitudes towards the SMS system, and how the SMS system fits into their daily life (Appendix V). The interviews were transcribed and later translated into English for analysis using a code book. These were back translated to ensure accurate translation.

Figure 8 Questions in the text message survey



Outcomes

The first outcome of the study was the complete survey response rate. Complete response was pre-defined as completing all questions in >70% of the daily surveys sent in a two-week round, or completing all questions in the fortnightly survey. The second outcome was the estimated rate of having diarrhoea during the two-week round, measured as having any number of loose or watery stools in the past two weeks. The third outcome was the attitudes towards the different surveying strategies, as uncovered in the qualitative work.

Quantitative Sample Size and Analysis

I conducted a simulation-based analysis of the design. I calculated delta, the minimal detectable treatment effect, for power of 80% and type I error rate of 5%. I assumed that the interaction effect sizes were half the size of the direct effect of each treatment. A baseline response rate of 50% was assumed as this was the most conservative value in terms of power. I also assumed an intra-class correlation coefficient of 0·05 for the proportion of variance at the individual level. Under these assumptions, the minimum detectable average treatment effect was 7·5 percentage points.

Complete response was analysed using a standard model for factorial trials[84]. A multivariable logistic regression model was estimated, containing indicators for each treatment and all treatment interactions, as well as demographics (participant sex, education, age, and household income) and random effects at the individual level. Dummy variables to adjust for time and area were also included. Average marginal treatment effects for each treatment in absolute terms (percentage point difference) were then estimated from each model.

The second outcome, diarrhoea rate, was examined in an identical way for complete responses – again with a logistic regression model, using the estimated

presence of loose or watery stool at any point during the two-week period as the dependent variable.

All data were monitored daily for any issues with receipt of the data. Due to the failure of the mobile phone network, rounds two and seven were repeated at the end of the survey period (disregarding any data from the first attempt of rounds two and seven).

Qualitative Sample Size and Analysis

Qualitative individual semi-structured interviews were conducted until theoretical saturation was reached. I determined that theoretical saturation was reached when both I and the research assistant conducting interviews did not find any themes reported by respondents in the previous two interviews.

Qualitative interviews were examined through the use of a codebook. I manually coded each interview by hand, looking for themes within broad topics (such as ease, questioning frequency, etc.). Topics of particular attention included time, convenience, comfort, and social harms. The codebook was then examined to determine the most frequent themes. Due to financial and human resource constraints, I coded the data alone.

Ethics Approval and Consent to Participate

The study was approved by the University of Warwick Biomedical and Scientific Research Ethics Committee (BSREC) (REGO-2018-2148), and cleared by local environmental health officers at the Mwanza Urban Water Supply and Sanitation Authority and the Lake Victoria Basin Office. I additionally followed international standards for health research as per University of Warwick and United Nations: participants were asked to give written informed consent; all documents were in the Tanzanian dialect of KiSwahili, and all documents were explained and read to the participant by a local research assistant if necessary. Data were stored on an

encrypted server. All participants were free not to take part; and in the event that a carer reported blood in the stool, were told to visit the public hospital. Further, all participants were given a local phone number to call with concerns. The protocol was prospectively registered on ISRCTN on the 20th of March 2019 under number ISRCTN11410773.

3.4.Results

Participant Demographics

In April 2019, one-hundred and forty-one respondents were recruited and randomised into one of eight arms. Figure 9 presents the CONSORT flow chart. There were no withdrawals, and all respondents who wished to take part were eligible. Table 6 reports the summary statistics for the study cohort. The average age of the respondents was 28·9 years – with most (92·2%) having completed primary school or above. Respondents were predominantly female (97·9%), and most (51·1%) had a household income below 50,000TZS (21·75USD) a month. The average household size was 5·5 people.

Figure 9 CONSORT Flow Chart

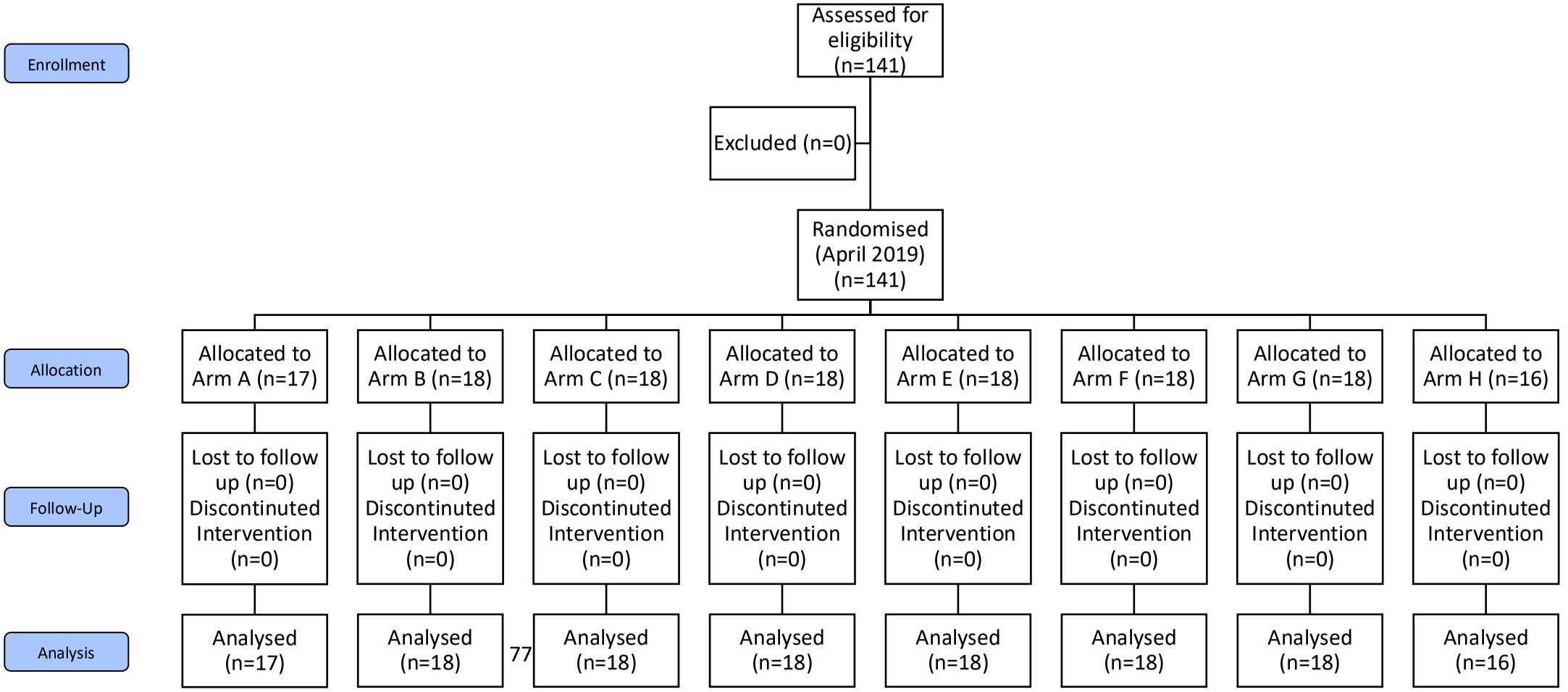


Table 6 Summary statistics of demographics and quantitative study outcomes for all participants, by study arm

		A	B	C	D	E	F	G	H	ALL
Total, n (%)		17 (12·1)	18 (12·8)	18 (12·8)	18 (12·8)	18 (12·8)	18 (12·8)	18 (12·8)	16 (11·3)	141 (100.0)
Age, Mean (SD)		28·9 (6·2)	28·4 (6·8)	27·6 (5·2)	30·4 (9·2)	30·3 (6·7)	29·3 (6·7)	29·9 (8·1)	27·9 (7·2)	29·1 (7·0)
Education, n (%)	None	0 (0.0)	0 (0.0)	0 (0.0)	1 (5·6)	0 (0.0)	0 (0.0)	2 (11·1)	1 (6·3)	4 (2·8)
	Some Primary	0 (0.0)	0 (0.0)	0 (0.0)	1 (5·6)	1 (5·6)	1 (5·6)	3 (16·7)	1 (6·3)	7 (5·0)
	Finished Primary	10 (58·9)	10 (55·6)	13 (72·2)	11 (61·1)	9 (50·0)	12 (66·6)	8 (44·4)	11 (68·8)	84 (59·6)
	Some Secondary	1 (5·9)	4 (22·2)	0 (0.0)	2 (11·1)	0 (0.0)	4 (22·2)	3 (16·7)	2 (12·5)	16 (11·3)
	Finished Secondary	6 (35·3)	4 (22·2)	3 (16·7)	2 (11·1)	8 (44·4)	1 (5·6)	2 (11·1)	0 (0.0)	26 (18·4)
	Some Tertiary	0 (0.0)	0 (0.0)	2 (11·1)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (6·3)	3 (2·1)
	Finished Tertiary	0 (0.0)	0 (0.0)	0 (0.0)	1 (5·6)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (0·7)
% Female, n (%)	Male	0 (0.0)	0 (0.0)	0 (0.0)	1 (5·6)	2 (11·1)	0 (0.0)	0 (0.0)	0 (0.0)	3 (2·1)

	Female	17 (100.0)	18 (100.0)	18 (100.0)	17 (94·4)	16 (88·9)	18 (100.0)	18 (100.0)	16 (100.0)	138 (97·9)
Income, n (%)	>50,000	8 (47·1)	10 (55·6)	8 (44·4)	12 (66.6)	8 (44·4)	9 (50.0)	8 (44·4)	9 (56·3)	72 (51·1)
	50,000-100,000	5 (29·4)	4 (22·2)	7 (38·9)	6 (35·3)	7 (38·9)	4 (22·2)	4 (22·2)	7 (43·8)	44 (31·2)
	100,000-500,000	2 (11·8)	3 (16·7)	2 (11·1)	0 (0.0)	3 (16·7)	3 (16·7)	4 (22·2)	0 (0.0)	17 (12·1)
	500,000+	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (5·6)	1 (5·6)	0 (0.0)	2 (1·4)
	Decline to say	2 (11·8)	1 (5·6)	1 (5·6)	0 (0·0)	0 (0·0)	1 (5·6)	1 (5·6)	0 (0·0)	6 (4·3)
Area, n (%)	Unguja	6 (35·3)	6 (35·3)	6 (35·3)	6 (35·3)	6 (35·3)	6 (35·3)	7 (38·9)	7 (43·8)	50 (35·5)
	Igogo	6 (35·3)	7 (38·9)	7 (38·9)	6 (35·3)	6 (35·3)	6 (35·3)	6 (35·3)	5 (31·25)	49 (34·8)
	Kilmahaewa	5 (29·4)	5 (27·8)	5 (27·8)	6 (35·3)	6 (35·3)	6 (35·3)	5 (27·8)	4 (25·9)	42 (29·8)
Household Size, Mean (SD)		5·6 (2·2)	4·9 (1·7)	5·7 (2·2)	5·4 (3·2)	5·1 (1·6)	5·9 (2·6)	6·8 (3·0)	4·8 (2·1)	5·5 (2·4)
14-day Diarrhoea Prevalence, %		34·7	27·1	38·3	42·9	40·6	30·4	24·0	52·4	36·4
Complete Response Rate, %		55·2	42·8	79·9	48·6	44·4	38·9	34·7	32·8	47·3

Survey Response Rates

Over the course of the study, between April and September 2019, 8,215 surveys were distributed: 7,655 daily texts and 560 fortnightly texts, with an even split between the high question load (3-question) and low question load (1-question) surveys, and incentive and no incentive. These can be broken down into 1,122 child-rounds of observation (each round lasting two weeks). The trial concluded in September 2019 when all arms progressed through all treatment combinations.

The mean response rate was 47%. Daily questioning had a similar mean response rate to fortnightly questioning (46·6 % vs 48·0%); the 3-question survey was lower than the 1-question (43·8% vs 51·0%); and the incentivised surveys was higher than the surveys without incentive (50·6% vs 44·0%) (Fig. 10). When examining mean response rates by interactions between treatments, there was little evidence of any interaction between treatments, other than response rates being lower when daily questioning and the 3-question survey were combined (Fig. 11). Response rates increased as the study progressed (Fig. 10).

Table 7 reports the results from the adjusted model-based analysis. Daily questioning was associated with a non-significant reduction in the response rate by a 1·2 percentage point difference (ppd) [95%CI:-4·9,2·5], compared to fortnightly questioning. The 3-question survey was associated with a significant reduction of response rates by 7·0ppd [-10·8,-3·1] compared to the 1-question survey. Incentivisation was associated with a significant increase in response rates by 6·5ppd [2·6,10·2] compared to no incentive.

There was also evidence that respondent age affected response rates, with each additional year of age being associated with an increased in response rate by 1·1ppd [0·2, 2·1], as did time, with a 0·9ppd [0·0,1·7] increase per round (Table 7). Having education beyond the primary stage was associated with an increase in

response rates by 11·7ppd [-1·6,25·1] when compared to having primary education or lower. Having a low income (below 50,000TZS) was associated with a decrease in response rate by 3·8ppd [-16·2,8·6] when compared to middle or high income (Table 7).

Table 7 Estimated Adjusted Treatment Effects and Effects of Demographic Factors on Response Rate

Treatment and Demographic Factors	**Adjusted Treatment Effect (percentage point difference, (95%CI))**
Daily Recall vs 14 day Recall	-1·2 (-4·9,2·5)
3-question Survey vs 1-question Survey	-7·0 (-10·8,-3·1)
Incentive vs No Incentive	6·4 (2·6,10·2)
Age of Respondent (continuous in years)	1·1 (0·2,2·1)
Beyond Primary Education vs Primary Education or Lower	11·7 (-1·6,25·1)
Low Income (below 50,000TZS) vs Middle or High Income	-3·8 (-16·2,8·6)
Study Round (continuous)	0·9 (0·0,1·7)
Kilimahewa vs Igogo	-3·4 (-18·6,11·2)
Unguja vs Igogo	-6·0 (-21·2,9·1)

Estimated Diarrhoea Rates

Overall, 36·4% of the 14-day child-rounds reported diarrhoea. When broken down by treatment, daily questioning had an estimated diarrhoea rate of 51·2% (compared to 21·9% for fortnightly questioning); the 3-question survey had a 36·3% estimated diarrhoea rate (compared to 36·4% for the 1-question survey); and the incentivised surveys had a 38·7% estimated diarrhoea rate (compared to 33·6% for surveys without incentivisation) (Fig. 10). When looking at the impact of interactions between interventions on diarrhoea rate, I see a similar trend, with all treatment combinations that included the fortnightly survey having a similar lower estimated rate, regardless of interaction (Fig. 11). The estimated diarrhoea rate appeared to decrease as the study progressed (Fig. 10)

Table 8 reports the results from the model-based analysis. Compared to fortnightly questioning, daily questioning was associated with a significant increase in the estimated diarrhoea rate, with an adjusted treatment effect of 29·9ppd [22·8,36·9]. There was no evidence to suggest that the 3-question survey had a significant impact on the estimated diarrhoea rate, with an adjusted treatment effect of 0·0ppd [-6·0,5·9]. There was little evidence indicating that financial incentivisation had a significant impact on the estimated diarrhoea rate, with the incentive raising the estimated diarrhoea rate by 3·0ppd [-3·1,9·0].

Evidence showed an impact by respondent age, with each additional year in age associated with a decrease in the estimated diarrhoea rate by 1·2ppd [-2·2,-0·2], but not by other demographics (Table 8). Evidence also indicated a decrease in the estimated diarrhoea rate over the course of the study by 2·9 ppd per round [-4·3,-1·5].

Table 8 Estimated Adjusted Treatment Effects and Effects of Demographic Factors on Estimated Diarrhoea Rate

Treatment and Demographic Factors	**Adjusted Treatment Effect (percentage point difference, (95%CI))**
Daily Recall vs 14-day Recall	29·9 (22·8,36·9)
3-question Survey vs 1-question Survey	-0·0 (-6·0,5·9)
Incentive vs No Incentive	3·0 (-3·1,9·0)
Age of Respondent (continuous in years)	-1·2 (-2·2,-0·2)
Beyond Primary Education vs Primary Education or lower	3·7 (-10·2,17·6)
Low Income vs Middle or High Income	-0·3 (-12·9,12·3)
Study Round (continuous)	-2·9 (-4·3,-1·5)
Kilimahewa vs Igogo	4·5 (-10·5,19·6)
Unguja vs Igogo	-1·6 (-16·7,13·6)

Figure 10: Mean response rates (1-3) and estimated diarrhoea rates (4-6) over all eight rounds to the SMS Survey, broken down by treatment

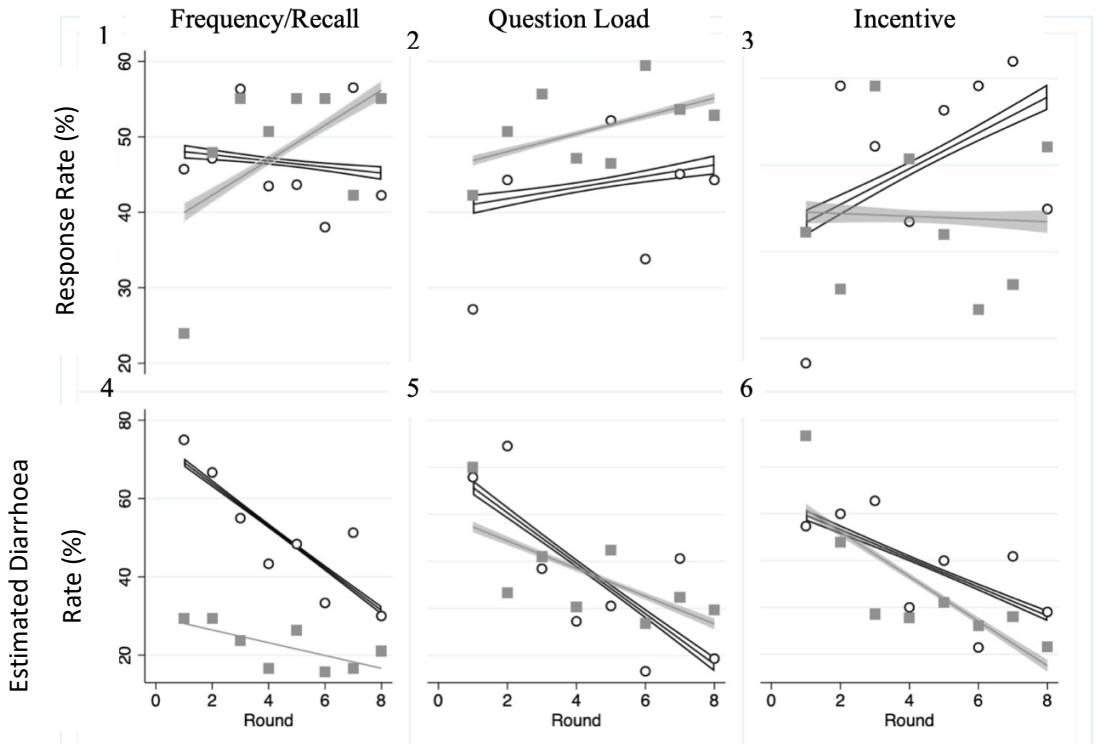


Figure 10 Caption: 1 and 4: daily (circle points) Vs Fortnightly (square points) surveys; 2 and 5: 3-Question (circle points) Vs 1-Question (square points) surveys; and 4 and 6: Incentive (circle points) Vs No Incentive (square points). Trend lines are displayed with corresponding 95%CI

Figure 11: Mean response rate (above) and estimated diarrhoea rate (below), and 95%CIs, of treatment combinations.

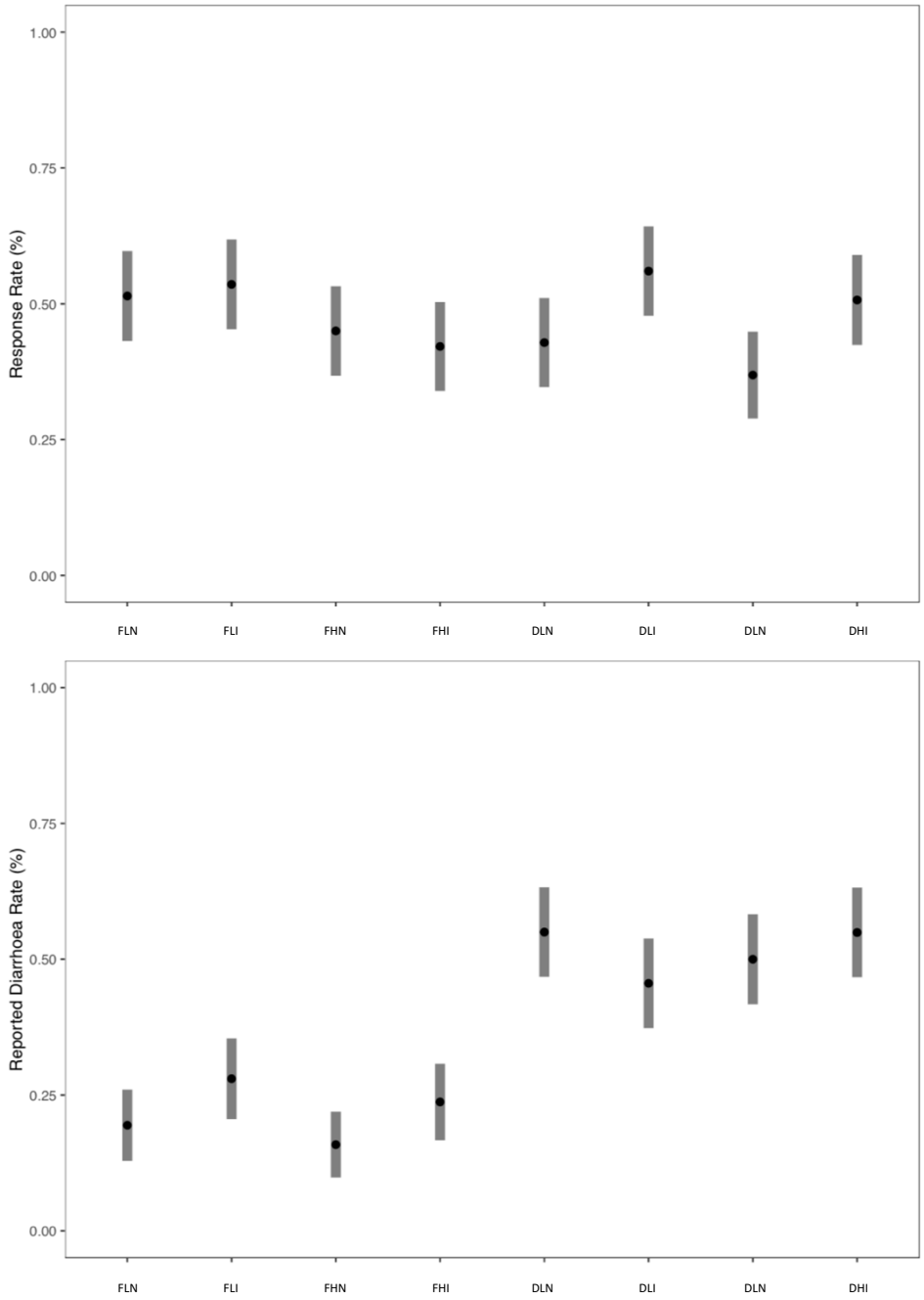


Figure 11 Caption: F: Fortnightly Questioning; D: Daily Questioning ; L: Low Question Load (1-Question Survey); H: High Question Load (3-Question Survey); N: No Incentive; I: Incentivisation

Qualitative Findings

Table 9 presents the key qualitative findings. A high degree of acceptance with the SMS surveying system was observed during the analysis of the qualitative interviews. This is likely because participants were accustomed to using mobile phones, as they use them in daily life for work, communicating with friends and family, and studying, saying, for example, “I use a mobile phone-especially communicating with customers and purchasing materials from suppliers;” and “'I am using my phone most of the times calling and answering calls and sometimes texting to my relatives.”

Participants reported that the messages were not perceived as intrusive and that late morning receipt of messages was convenient. Participants further reported appreciation of the reminders, with one reporting “The reminders offered was good especially for question one, I always forget to answer but the reminders assisted me to answer the questions.”

Participants were mixed regarding if incentivisation altered how likely they were to respond. Some stated that incentivisation “encouraged [them] a lot,” as it provided them with the opportunity to call family. Others said it did not factor into if they replied. Participants also reported appreciation for being able to report on their child’s health, and would prefer if the study went further – checking on the child after health conditions were reported, offering education, and provided money for food or medicine. One participant stated “The project should be able to offer education on the type of food to be given to children especially after treatment. The project should also check on children after they discover an issue and provide money to buy food and medicine for the children.” Participants also reported paying more attention to their child’s health because of the survey.

Participants generally did not mind frequent questioning, but believed they would be able to remember diarrhoea that occurred 14 days ago. Participants also did not mind answering more questions, with some even wanting to answer more questions on their child health.

Participants were mixed regarding preference towards face to face surveys vs SMS surveys. Some preferred the ease and privacy of SMS surveys, with one stating “I would be able to talk to a person but I prefer the privacy by phone. Others, however, preferred face to face surveys, as 1) they provide the opportunity to ask questions; and 2) do not face pressure from other family members.

Table 9 Key Qualitative Findings

Broad Theme	Code	Example 1	Example 2
Phone Use in Life	**Social Life**	I am using my phone most of the times calling and answering calls and sometimes texting to my relatives	
	Work Life	I use a mobile phone-especially communicating with customers and purchasing materials from suppliers,	
Convenience	**Initial Message Convenient**	They were not annoying or disturbing my timetable, the timing were very reasonable though sometimes they came when I am on home duties	Messages which I received they were coming in the morning hours so it was convenient time for me
	Reminders Useful	The reminders offered was good especially for question one, I always forget to answer but the reminders assisted me to answer the questions	
Motivation/Incentivisation	**Behavioural Nudge for Health**	Also I can see the difference in my attention especially to the stool of my child; I received one SMS telling me to go to the hospital because was happy receiving those instructions hence went to a District hospital in Butimba.	The project should be able to offer education on the type of food to be given to children especially after treatment. The project should also check on children after they discover an issue, and provide money to buy food and medicine for the children
	Incentive - needed	The amount of money encouraged me a lot	I had to send a reply because the money sent to me helped me even to call my relatives

	Incentive - not needed	If there could be no money I still could have responded to the messages.	Incentive offered did not contribute to me feeling pressured to give answers
Ease of use	**Easy**	I never faced any obstacle and the questions were fair and understandable.	Technology was not a problem at all I was able to respond to all questions easily
Question Frequency	**Did not mind high frequency**	I just receive two questions every day for about a month and because it is for a child only I think those two questions are enough,	
	Thought could remember past 14 days	I felt comfortable to answer questions which she will be able to remember for the questions asked two weeks ago because these are the information of her child but it is not easy to remember information for the children of other people	
Question Levels	**Preferred more questions**	Questions should be increased	
SMS v In Person	**SMS preferred**	A message is good because meeting me sometimes is difficult because I might be at work in a far place	I would be able to talk to a person but I prefer the privacy by phone
	Person preferred	Talking to a person direct is easier and better because there is a chance to ask questions in case you don't understand or if you have more	I feel comfortable to answer the survey; I prefer face to face rather than using phone no friend, family or community who force/give out pressure to complete answering questions/survey

3.5.Discussion

I conducted an individual level factorial multiple crossover randomised control trial in Mwanza, Tanzania to estimate the effects of questioning frequency, question load, and incentivisation on response rates to an SMS survey on under-five diarrhoea in urban informal settlements. The study also included analyses of the effects of demographics on the response rate; the effects of questioning frequency, question load, and incentivisation on the estimated diarrhoea rate; and a qualitative examination of attitudes towards text message surveys. The principal findings of the study are that SMS messaging can be a suitable means of disease surveillance in LMICs, with response rates of around 50% in my study, but that results can be impacted by the methodologies used: financial incentivisation is associated with an increase in the response rate, increased questioning loads is associated with a decrease in the response rate, and frequent questioning with short recall periods is associated with a decrease in the estimated diarrhoea rate.

The Impact of Treatments on the Response Rate

The complete response rate over all eight rounds was 47% - a proportion higher than reported in previous similar studies: 15% in Liberia during the Ebola Outbreak, and 31% in Ghana during a demographic and health survey[44, 45]. Reasons for the higher response rate include differences between the study sites, health topic, and study recruitment (with my study recruiting consenting participants, whereas the aforementioned studies randomly messaged unconsented participants). Evidence from my study additionally indicates that daily questioning (24-hour recall) had a similar response rate to the fortnightly survey (14-day recall), suggesting that after 14 days of daily questioning respondent fatigue did not set in, as has been suggested in past studies[85]. Further supporting that fatigue did not set in, there was an increase in the response rate over time. Additionally, qualitative work revealed that participants did not mind answering questions more frequently.

There did appear to be a slightly lower response rate for the 3-question survey (when compared to the 1-question survey) – suggesting that the 3-question survey was burdensome for the respondents. This is in line with past studies, including Bhavnani et al. (2014), who suggested that fatigue might occur if participation required a high amount of effort[85]. Respondent fatigue is supported by my finding that when daily questioning is combined with the 3-question survey, there is an additional lowering in response rate. This reduction in response rate may have been even more apparent if regardless of response participants were given all three questions, rather than only being given all three questions if the participant reported diarrhoea. This was not supported by qualitative findings, however. Participants generally reported wanting to answer even more questions than we asked.

The incentive did yield a statistically significant increase in the mean response rate. This is consistent with Hopkins and Gullickson's (1992) meta-analysis on the impact of financial incentivisation on survey response, which found that financial incentivisation increased response rate by 19% when given with the survey (prior to completion) and by 7% when given after the survey. The latter figure is similar to the 6·4ppd increase in response rate observed in my survey through provision of an incentive after survey completion[86]. Qualitative results were mixed regarding the impact of incentivisation – with some participants saying it pushed them to respond, with others saying it did not matter. However, the participants stated that non-financial incentivisation may also work, including: 1) providing health education with the messages; 2) checking on children who are reported as unwell; and 3) providing medicine or food for the children. The participants also stated that the messages encouraged them to pay more attention to the health of their children, which was a large motivator.

The Impact of Demographics on the Response Rate, and Ease of Use

Those with higher education were more likely to respond, with those who had progressed beyond the primary stage of education responding at a rate 11·7ppd higher than those with primary education or below. This finding was also seen in L'Engle et al. (2018) study on demographic and health surveys Ghana[45]. L'Engle et al. surveyed a nationally representative sample using an eighteen question demographic and health survey to determine response rates to a mobile phone survey[45]. This study used pre-recorded voice messages in which participants would respond by inputting a certain number on their dial pad[45]. Comparing the results of the mobile phone survey to two similar nationwide surveys which used face to face surveying, the study estimated that populations with no education answered the mobile phone survey at a rate 5 to 18 ppd less than a face to face survey[45]. The study also estimated that populations with secondary education or above answered the mobile phone survey at a rate 27 to 29 ppd higher than a face to face survey[45]. L'Engle et al. concluded that while mobile phone surveys are a promising tool for data collection, differential response rates by varying demographics could introduce bias if adjustments were not made.

Qualitative work revealed that participants generally did not have issues responding to the survey. The messages generally came at a convenient time, and if they did not the reminders were appreciated. Participants also reported ease in using the SMS survey system, as they often used their mobile phones in daily life for work and socialising. This is unsurprising, considering how common mobile phones are in urban East Africa[59].

The Impact of Treatments on the Estimated Diarrhoea Rate

Diarrhoea was reported in 36% of complete child-rounds – yielding an estimated incidence of 9 episodes per child year. While this number is slightly higher than previously estimated in urban East Africa, the finding can be explained on the basis

that all participants in the previous studies were presented a 14-day recall period[87]. When restricting the analysis to the 14-day recall period, I estimated an incidence of 6 episodes per child year, in line with previous studies. This is considerably lower than the estimated incidence of 13 episodes per child-year for 24-hour recall. The higher diarrhoea rate estimated for the daily survey with 24-hour recall, when compared to the fortnightly survey with 14-day recall, provides support of recall bias, whereby respondents forget events that occur over long periods[33, 62]. Feiken et al. (2010) report prevalence dropping from 18% for 24-48 hour recall to around 5% in 11-13 day recall[33, 62]. Zafar et al. (2010) report that severe diarrhoea is twice as likely to be reported as moderate diarrhoea during longer recall periods[33, 62]. This conflicts with qualitative findings in which parents report being able to recall diarrhoea over 14 days.

Incentive and survey type did not influence estimated diarrhoea rates. Of interest, however, estimated diarrhoea rates did decrease markedly over subsequent rounds. I hypothesize three (non-exclusive) reasons for this. First, the survey may have created a heightened awareness of diarrhoea risk and child health (as reported in the qualitative work), resulting in better WASH practices; second, respondents may have been embarrassed by constantly reporting diarrhoea[88]; third, respondents may have telescoped answers at the beginning of the survey (recalling from a longer period than the stated recall period).

Strengths and Weaknesses

There are two substantial weaknesses in this study. 1) As the questioning frequency treatment included variation of both frequency and recall period – with fortnightly questioning asking about the past fortnight, and daily questioning asking about the past day, it is not possible to determine if the differences associated with this particular treatment were due to the frequency or the period of recall. 2) The study was unable to ascertain the impact of perception bias and if participants truly

understand what defines a case of diarrhoea. For example, in a previous study, Voskuijl et al. (2017) found that carers of infants with severe acute malnutrition in a Malawian hospital were only able to identify 75% of loose or watery stools as such (loose or watery stools being identified by observation by a health care provider)[89].

This study has several strengths. The study took place in an urban East African city with a fairly representative culture and geography of other urban East African areas, so I believe that the results are generalisable to similar settings. Further, the study data provides results which are not only consistent throughout the study, but also build of past literature. Bhavnani et al. (2014) discussed the possibility of respondent fatigue through frequent, in depth, questioning which I provide evidence for[85]. Hopkins and Gullickson (1992) found evidence that incentivisation is associated with increased response rates, which I also find evidence for, but in the novel form of an SMS survey[86]. Similarly, Feiken et al. (2010) and Zafar et al. (2010) found evidence for recall bias during in-person surveys for diarrhoea, which I also see in my novel SMS survey[33, 62]. Finally, L'Engle et al. (2018) found a substantial association of demographics, such as education, on response rate in their SMS survey, but, due to their use of uninformed participants, had a low response rate[45]. My use of informed participants resulted in a higher response rate.

Conclusion

SMS surveying is a feasible method of collecting data on child health among populations with high levels of access to mobile phones. There are several variations in the system which may affect response and estimated diarrhoea rates. Financial incentivisation (compared to no financial incentivisation) increases the response rate but does not impact the estimated diarrhoea rate. A high question load (3-questions compared to 1-question) decreases the response rate, particularly when done so at a high frequency, but does not impact the estimated diarrhoea rate. Daily questioning and recall (compared to 14-day questioning and

recall) does not impact response rate, but dramatically increases the estimated diarrhoea rate.

When conducting standard in-person active surveillance, my results call for the need to standardise the methodologies used to minimise undesirable variation in results. These standardised methodologies should use incentivisation and low question loads to maximise response rate, while using a short recall period to minimise recall bias.

Future research is needed in this field, however, including evaluation of SMS surveillance systems in other populations, such as those in rural areas; evaluation of questioning frequency and recall period separately; and further evaluation into the ability of carers to correctly identify diarrhoea. In the next chapter I look at the ability of carers to correctly identify diarrhoea, as well as the extent of misclassification when using diarrhoea as a proxy of enteric infection.

4. Evaluating Diarrhoea as a Proxy Marker of Enteric Infection: A Comparison of Traditional Diarrhoea Measurement Methods with Microbiological and Biochemical Indicators in the Cox's Bazar Forcibly Displaced Myanmar Nationals Camp: A Cross-Sectional Observational Study

4.1.**Abstract**

Introduction

Water, Sanitation, and Hygiene (WASH) systems aim to reduce enteric infection, particularly among children under the age of five. The most often used primary outcome of WASH trials is carer-reported diarrhoea. To evaluate intervention effectiveness using this primary outcome, it is therefore necessary to assume that diarrhoea classifies enteric infection reliably. In this chapter, I evaluate diarrhoea as a proxy marker of enteric infection.

Methods

I performed a study in the Cox's Bazar Forcibly Displaced Myanmar Nationals Camp. First, I randomized 800 households into either a standard questionnaire asking carers if their under-five has had diarrhoea in the past fortnight; or a pictorial questionnaire asking carers to pick from a stool chart which stools their under-five has had in the past fortnight. Second, I collected stools from a random sample of 120 of the under-fives. Stools were examined visually, tested for proteins associated with enteric infection, and tested for 26 common enteric pathogens. Using the enteric pathogen tests as a 'gold standard,' I calculated sensitivities and specificities for each survey type, visual examination, and proteins as proxy markers of enteric infection.

Results

The probability of a carer whose child's stools had at least one enteric pathogen detected reporting diarrhoea (sensitivity) using the standard survey was 0.49 [95%CI:0.32 ,0.66] and the probability of a child whose stools did not have any enteric pathogens detected not reporting diarrhoea (specificity) was 0.65 [0.41, 0.85]. Qualitatively similar sensitivities and specificities were observed for the other proxy markers.

Implications

While diarrhoea is an important sign in clinical practice it appears that it is a poor proxy for enteric infection in epidemiological surveys in Cox's Bazar and likely other similar environments. In WASH trials, intervention effects are therefore likely to be biased towards the null.

4.2.Introduction

Despite WASH interventions aiming to disrupt the transmission of enteric pathogens, trials of WASH interventions generally do not use microbiologically confirmed enteric infection as their primary endpoint. As diarrhoea is the main clinical presentation of enteric infection at health facilities, as well as an important cause of child death, diarrhoea is typically used as the end point in WASH trials. Diarrhoea is also cheaper to measure than enteric infection. Given that WASH interventions only affect the risk of diarrhoea indirectly, via a reduction in the risk of infection, we must assume that diarrhoea can classify enteric infection reliably to reliably estimate intervention effectiveness. However, many recent large-scale evaluations of WASH interventions have found limited evidence for intervention effectiveness using diarrhoea as an endpoint[18-20]. It has been hypothesized that the lack of intervention effectiveness is due to diarrhoea not classifying enteric infection reliably.

WASH trials generally elect to use diarrhoea rates estimated from active surveillance techniques (surveying in the community) rather than passive surveillance techniques (measuring health facility visits for diarrhoea). The use of active surveillance is due to many children with diarrhoea not visiting health facilities. For every case of diarrhoea that presents at a health facility there will be about 50 reported on a community based survey[90]. Active surveillance techniques therefore likely estimate much more accurate diarrhoea rates than passive surveillance, as we have demonstrated in Chapter Two. However, active surveillance can also introduce bias and error into estimated diarrhoea rates, as seen in Chapters Two and Three.

From a clinical point of view benign carriage of an infective organism is of little interest; it is well known that an infected person can be entirely well. But this is not true from an epidemiological or public health perspective; a carrier can still spread

the organism and carriage is therefore an indication of population risk. It is this risk that WASH interventions aim to reduce. So, in this study I set out to measure the extent of misclassification when using diarrhoea as a proxy marker of infection. Until recently direct measurement of the carriage of pathogenic organisms has been difficult due to high cost and the need for costly infrastructure that is often unavailable. However, with falling costs of stool pathogen testing, widespread availability of lab equipment, and the advent of rapid diagnostic tests (e.g. lateral flow tests using immunochromatography) with good sensitivities and specificities for infection, measurement of enteric infection is possible.

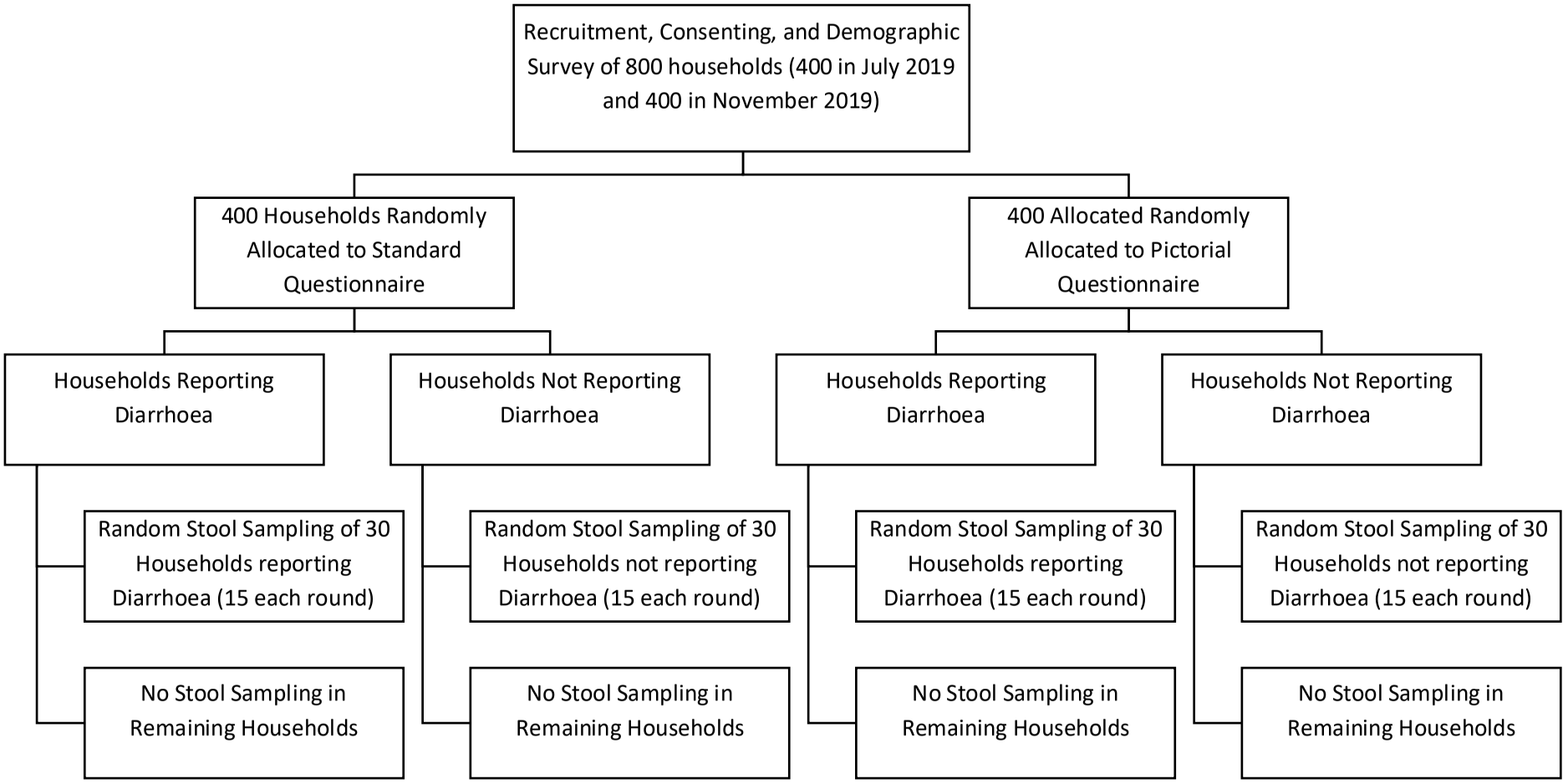
To determine the extent of misclassification when using diarrhoea as a proxy marker of enteric infection, I conducted an observational study in the Cox's Bazar Forcibly Displaced Myanmar Nationals Camp. I evaluated four methods of measuring diarrhoea as proxy markers of enteric infection: 1) the standard survey method, recommended by UNICEF and the Demographic and Health (DHS) survey, that is used in most evaluations of WASH interventions: asking carers if their child has had diarrhoea in the past 14 days[63]; 2) pictorial surveys, asking carers to pick from the Amsterdam Stool Chart which stools their child has had in the past 14 days[91]; 3) visual confirmation, where a trained researcher looks at the stool and determines if it is loose or watery[37]; and 4) analysis of the stool for proteins associated with diarrhoea (calprotectin and lactoferrin)[50, 51]. In addition to determining the extent of misclassification of enteric infection when using each of these proxy markers, I also calculated the agreement between standard verbal surveying, pictorial surveying, and visual stool analysis in measuring diarrhoea.

4.3.Methods

Study Design

I conducted a repeated cross-sectional randomized study in the Leda Makeshift Camp of Cox's Bazar at two time periods, the wet season in July/August 2019, and the dry season in November 2019. I aimed to estimate the misclassification error associated with four proxy markers of enteric infection in children under-five. I randomised 800 participants in a 1:1 ratio to either a standard or pictorial questionnaire on diarrhoea, described below. I then randomly collected stool samples from 15 participants not reporting diarrhoea and 15 participants reporting diarrhoea for each survey type in each of the two rounds for a total of 120 stool samples (Fig. 12). Stools were examined visually, and tested for proteins associated with enteric infections and 16 common enteric pathogens. A random sequence dictating from whom to collect stool was generated using Microsoft Excel, taking into account expected rates of reported diarrhoea to ensure that samples were collected evenly over time.

Figure 12 Study Design



Study Setting

Cox's Bazar, an area on the South-Eastern Coast of Bangladesh, is home to over one million Forcibly Displaced Myanmar Nationals (FDMNs). This population, mostly belonging to the Rohingya ethnic minority group, came to Bangladesh in several waves since 1962, with the most recent wave in 2017. Cox's Bazar contains 29 individual FDMN camps. This study took place in the makeshift camp area of Camp 24 (Leda), an area of high population density and poor WASH infrastructure, including poor drainage and a lack of sanitation facilities (Fig. 13). The first round of data collection took place in the time period between July and August 2019, the rainy and warm season (average temperature 32C with 25-27 rainy days per month) in Cox's Bazar; and the second round of data collection took place in the time period between November and December 2019, the dry and cold season (average temperature 28C with 2-3 rainy days per month) in Cox's Bazar[92].

Figure 13a-c Map of Bangladesh, Cox's Bazar, and Leda Makeshift Camp

Figure 13a Caption: A map of Bangladesh with the Cox's Bazar Region highlighted in the red box (Google Maps, 2021)

Figure 13b Caption: A map of the Cox's Bazar Refugee Camps, with Leda highlighted in the red box (Azad et al., 2019)

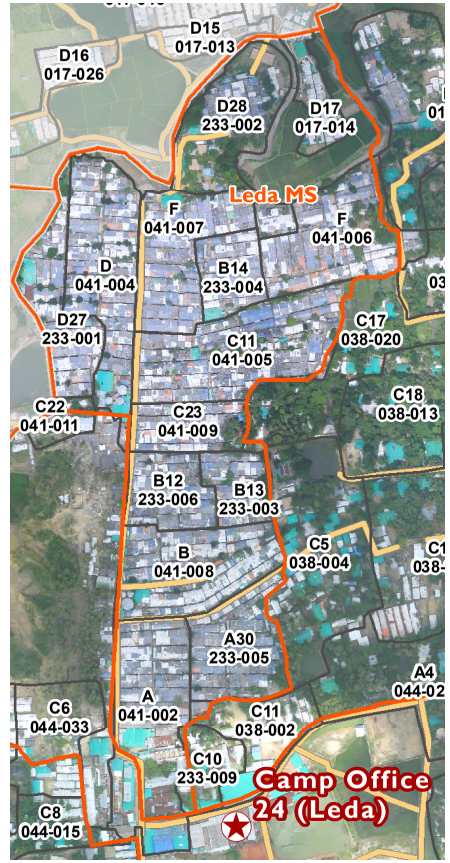


Figure 13c Caption: A drone overhead map of the Leda Makeshift Camp (ISCG, 2019)

Participant Recruitment and Randomisation

I aimed to recruit a random sample of 400 households during each round of data collection. As I could not obtain a community list of households, I used a variant of the extended program on immunization (EPI) sampling method, in which a series of random angles and distances from a predetermined central point were generated, selecting households closest to the random point[93]. Prior to the start of the day, a list of houses was generated using the EPI sampling method, with an arm randomly assigned using the RAND function in Microsoft Excel. If a selected household did not have an adult present, I returned daily until I made contact. If the household reported not having a child under the age of five, I moved onto the next house. Adults in households with a child under the age of five had the study explained to them. If the respondent was interested, I sought consent and evaluated the household against the inclusion and exclusion criteria: having at least one child under the age of five, the respondent being over the age of 18, and the household not expecting to relocate, resettle, or repatriate in the next six months (as to not preclude the possibility of follow up). The enrolment of consenting households included collection of the household's GPS coordinates, the carer's name, and the carer's mobile phone number. These identifiable data were recorded separately from the survey data and linked through individual household identifiers. In the event that a household had more than one child under five, I selected the oldest child under five as the main child of interest.

Data Collection and Variables

Demographic Survey

The survey began with a demographic survey, using questions extracted from the Demographic and Health Surveys[94]. The demographic survey included questions on the health and demographics of the carers and all children under five in the household, including: age, education, employment, health conditions, and time

spent in Bangladesh. I also asked about: the number of other people living in the household; household access to and use of WASH resources; building materials of the home; and breastfeeding. A full questionnaire can be found in Appendix VI. After this demographic survey, either the standard or pictorial diarrhoea questionnaire was administered.

Diarrhoea Questionnaires

In the standard questionnaire arm, I used the UNICEF/DHS recommended method of asking carers if their oldest child under five has had diarrhoea in the past 14 days (defined as three or more loose or watery stools in one day); if their child has had blood in their stool in the past 14 days; and, only if the carer reported diarrhoea or blood in stool, if and where they had sought care[63].

In the pictorial questionnaire arm, I measured the mid-upper arm circumference (MUAC) of oldest child under five. I then showed the respondents the Amsterdam Stool Chart (Fig. 14) and asked which consistencies and colours of stool their oldest child under five had in the past 14 days[95]. If the carer reported diarrhoea, which I defined as type A (watery stool), or type B (loose stool), I then asked how many times the child had this stool type on the worst day, how long ago they last had this stool type, and how many days the diarrhoea lasted. I then, regardless of diarrhoea status, asked about fever, blood in stool, vomiting, not eating, rashes, and eye problems in their infant over the last 14 days. If respondents answered yes to any of these, I asked if and where care was accessed.

Figure 14 Amsterdam Stool Chart

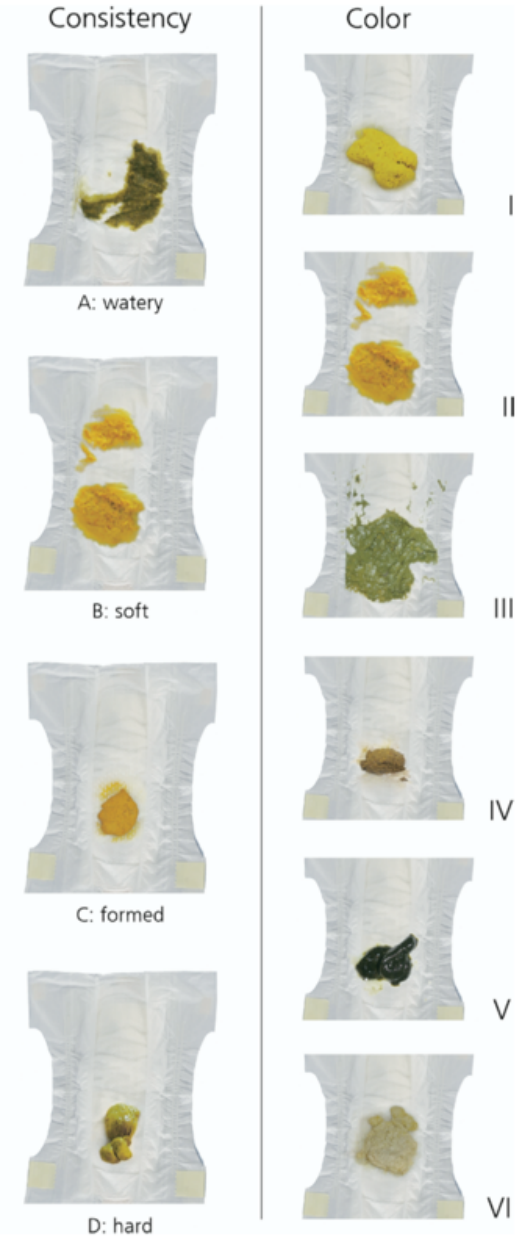


Figure 14 Caption: Amsterdam Stool Chart, with showing stools in diapers with Consistency A (Watery) to D (Hard), and colour I (yellow) to VI (pale gray) (From Bekkali et al., 2009)

Stool Sampling

I asked a random subset of carers to provide stool samples from their oldest child under-five. I gave those respondents who agreed a large container and instructed them that their oldest child under five should defecate into it, and that I would return the next morning to collect it. I then visited these households the next morning to collect the stool. If the household did not provide stool, I asked the reason why and if they would be able to provide a sample the next day. If they did not agree to provide stool, the household was marked as lost to follow up. Further, if the household was not reachable after two days, or did not provide a stool sample by Thursday of the week (the last day before the weekend), they were marked as lost to follow up. Replacements were sought through the collection of additional samples from a top up sample. When a household provided a stool sample, it was examined by a trained researcher to determine if it was loose or watery, and then transferred the into a specimen container and placed in an ice box. Within 8 hours, the stool was frozen to -20C and later transported to Dhaka for laboratory analysis. Stool was tested by Enzyme Linked Immunosorbent Assay (ELISA) for Calprotectin and Lactoferrin, and by qualitative Polymerase Chain Reaction (PCR) for 16 common enteric pathogens using techniques as recommended by the Centers for Disease Control[96]. The pathogens which I intended to test for can be found in Table 10.

Impact of the COVID-19 Pandemic

As described in the above, we had intended to test the stools for 26 endemic pathogens. While fieldwork was completed shortly before the COVID-19 pandemic, stool testing was due to begin just before the pandemic. Due to constraints on the laboratory resulting from lockdowns, staff illness, and prioritization of laboratory equipment for COVID-19 testing, not all pathogens were tested for. In total, we tested for 16/25 pathogens (Table 10). I do believe, however, that I have captured

the most important pathogens in these 16 – with the exception of *Ascaris Lumbricoides*.

Table 10 Pathogens Tested

Type	Species	Pathogens Not Tested Due to COVID-19
Bacteria	Shiga toxin-producing Escherichia coli	Tested
	Enteropathogenic Escherichia coli	Tested
	Enterotoxigenic Escherichia Coli	Tested
	Enteroinvasive Escherichia Coli	Not Tested
	Enterohemorrhagic Escherichia Coli	Not Tested
	Salmonella Species	Tested
	Campylobacter	Tested
	Shigella	Tested
	Cholera	Tested
	C. difficile	Not Tested
	C. perfringens	Not Tested
	Acinetobacter	Not Tested
Protozoa	Entamoeba histolytica	Tested
	Cryptosporidium	Tested
	Cyclospora	Tested
	Giardia	Tested
Helminth	Ancylostoma duodenale	Not Tested
	Ascaris lumbricoides	Not Tested
Viruses	Rota-virus	Tested
	Hepatitis A	Not Tested
	Hepatitis E	Tested
	Norovirus	Tested
	Human Caliciviruses	Not Tested
	Adenovirus	Tested
	Astrovirus	Tested
Fungi	Enterocytozoon	Not Tested

Statistical Methods

Sample Size Estimation

The sample sizes used (400 in the standard survey and 400 in the pictorial survey, split evenly between each of the two rounds) provided power to estimate the carer-reported diarrhoea rate in each arm (in both rounds combined) to a 95% confidence interval of approximately ±4 percentage points, assuming a point prevalence of 10%. The subsample providing stool samples (n=120) would provide precision to estimate the sensitivity and specificity values to approximately ±9 percentage points assuming values of 90%. Between the pictorial and standard questionnaire, using a 95% confidence interval will allow for a minimum detectable difference in reported diarrhoea rates of approximately 8 percentage points.

Data Analysis

I summarised the survey data by calculating the means and 95% confidence intervals of key demographic variables and risk factors for diarrhoea. These were broken down by round, data collection arm, and stool collection status. I similarly summarized diarrhoea results from each measure of diarrhoea in each round, with 95% confidence intervals. I defined 'having diarrhoea' for each measure as: 1) answering yes to the standard survey; 2) stating diarrhoea types A or B on the pictorial survey; 3) having faecal calprotectin level over 50ug/g and/or a faecal lactoferrin level over 7.25ug/g for protein measurements[50, 51]; or 4) stool visually being loose or watery for visual confirmation. I summarized results of the stool tests for enteric infection, broken down by survey type, round, and diarrhoea status – with 95% confidence intervals calculated. I additionally calculated overall rates of each enteric infection in each round as well as rates of having at least one enteric infection, referred to as "any infection", weighted by data collection arm and diarrhoea status.

I calculated the sensitivities and specificities for each of the four measures of diarrhoea as proxy markers of enteric pathogens being detected in stool. I did not calculate positive predictive value (PPV) and negative predictive value (NPV) as I artificially manipulated the underlying diarrhoea rate in the sample through the 50/50 sampling design. As such, estimates of PPV and NPV would be misleading. The 'gold standard' measure used was PCR detected infection (described above), compared against the proxy reporting 'having diarrhoea' as defined above. For each I could then determine the proportion of proxy marker results that are true positive (infectious stools with diarrhoea), false positive (non-infectious stools with diarrhoea), true negative (non-infectious stools without diarrhoea), and false negative (infectious stools without diarrhoea). These are presented in Table 11. Table 11 also presents the calculations I performed to calculate sensitivity and specificity (using the *diagt* function in StataSE Version 15). The same function also estimated standard errors and 95% confidence intervals[97]. These diagnostic performance indicators were calculated for infection as a whole (the gold standard being PCR detected any infection), and for specific categories of infection (e.g. PCR detected any bacterial infection). I additionally plotted the sensitivities and specificities of the aforementioned markers of any enteric infection on a ROC space plot. Finally, the agreements between the different measures for diarrhoea were determined through computation of their kappa values, along with expected and observed agreements.

Interpretation of Proxy marker Performance Indicators

Sensitivity is the probability of the proxy marker reporting the patient 'having diarrhoea' when the patient has an infection: for example, if 90% of children with an enteric infection 'had diarrhoea', the proxy marker would have a 90% sensitivity. Specificity is the opposite, the probability of the proxy marker reporting 'not having diarrhoea' when the patient does not have an infection: if 90% of children without an enteric infection did not 'have diarrhoea', the proxy marker would have a 90%

specificity. A proxy marker which tends to overestimate infection rates would likely have a high sensitivity and low specificity, with the opposite true for a proxy marker which would estimate lower levels of diarrhoea. When using the proxy markers of enteric infection to measure the relative risk of enteric infection during trials of WASH interventions, low specificities (even when equal in the control and intervention arms) would bias results towards the null, regardless of any true intervention effect[81].

Table 11 Calculations of Sensitivity, Specificity, PPV, and NPV

	Enteric Infection Present	**Enteric Infection Absent**	***Row Total***
Proxy marker Positive	• Infectious Stools • Diarrhoea	• Non-Infectious Stools • Diarrhoea	*All Diarrhoea Stools*
Proxy marker Negative	• Infectious Stools • No Diarrhoea	• Non-Infectious Stools • No Diarrhoea	*All No Diarrhoea Stools*
Column Total	*All Infectious Stools*	*All Non-Infectious Stools*	
Sensitivity=Infectious Diarrhoea Stools/All Infectious Stools Specificity=Non-Infectious Non-Diarrhoea Stools/All Non-Infectious Stools			

Ethics

Ethics was granted by the University of Warwick Biomedical Science Research Ethics Committee in the United Kingdom (REGO-2019-2345) and by the ICDDR,B Ethical Review Committee in Bangladesh (PR-19027). The study was also approved by the

Refugee Relief and Repatriation Commission, Government of Bangladesh (Letter 789) and Camp 24 authorities.

Full written consent was obtained, with a full data management and data collection plan in place. No identifiable data were shared with government authorities, UN agencies, or NGOs. Any participants who showed signs of dehydration or malnutrition were referred to a local clinic for treatment; and any safeguarding concerns were reported to the ICDDR,B head of mission for escalation as appropriate. No compensation was offered. The study was registered on ISCRTN during its inception, and updated as necessary (ISRCTN41564300).

4.4.Results

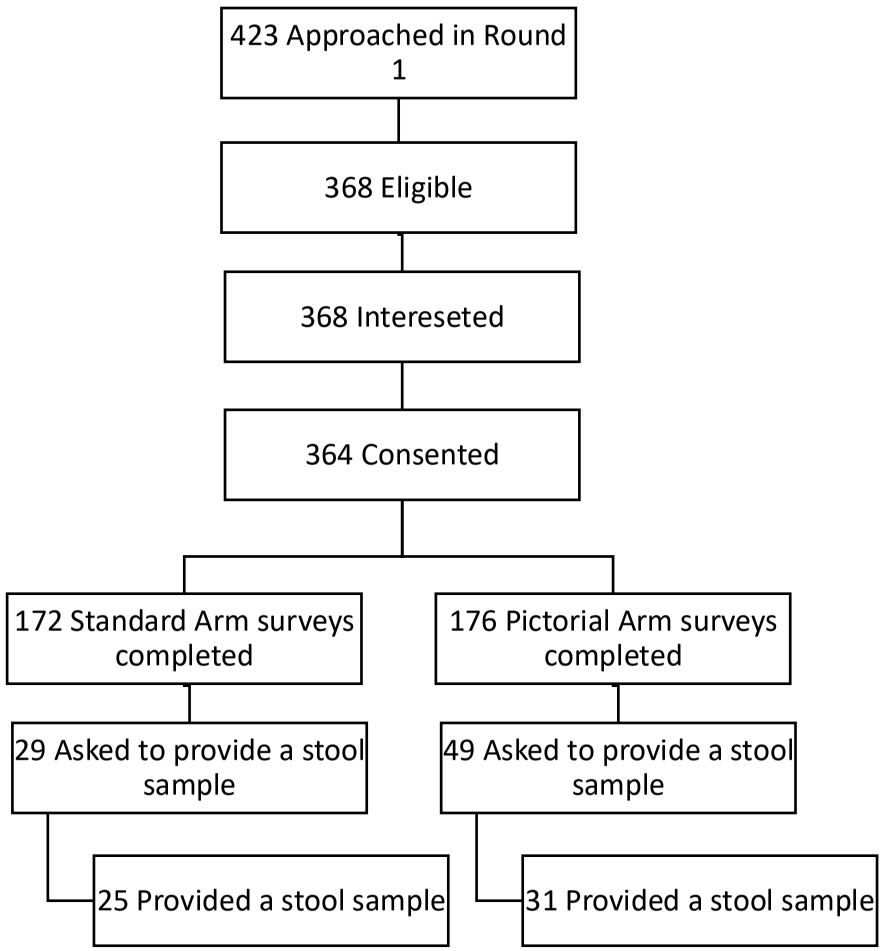
Study Recruitment

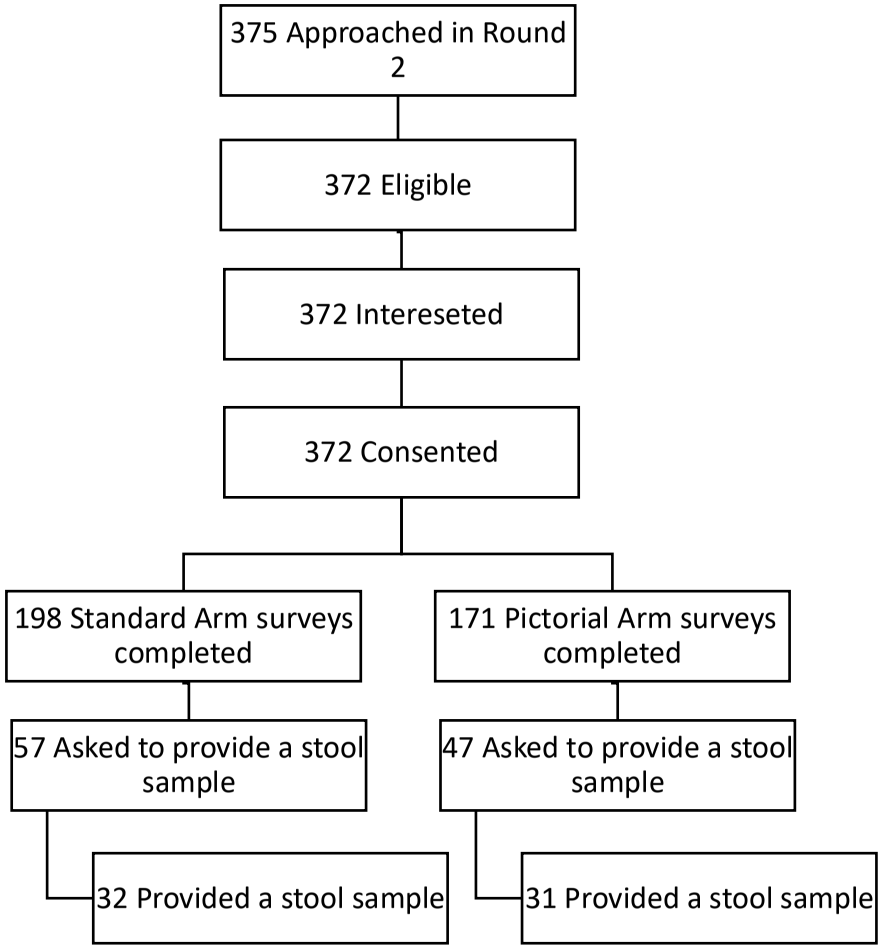
Figure 15 shows the participant flow chart. In round one, represented by the first column of the flow chart, I approached 423 households. These included 368 eligible households. All but four of these eligible households consented to take part in the study, of which 348 completed the survey (172 completing the standard survey, and 176 completing the pictorial survey). The sixteen that did not complete the study either withdrew consent during data collection, or had to leave during data collection (e.g. to collect water). I asked 78 of the households that completed the study to provide a stool sample and 56 households did so (25 in the standard arm, and 31 in the pictorial arm).

In round two, I approached 375 households. These included 372 eligible households. All eligible households consented to take part in the study, of which 369 completed the survey (198 in the standard arm and 171 in the pictorial arm). The five which withdrew, as in round one, either withdrew consent during data collection or had to leave for another reason. I asked 104 of these households to provide stool samples, of which 63 provided samples (32 in the standard arm and 31 in the pictorial arm).

Due to civil unrest in round one and Cyclone Bulbul in round two, I concluded data collection early in both rounds. This resulted in a shortfall of 52 surveys completed (348/400) and 4 stool samples collected (56/60) in the first round, and 31 surveys completed (369/400) in the second round. Sixty-three planned stool samples were collected in the second round, exceeding the planned number.

Figure 15 Recruitment Flow Chart





Summary Statistics

Table 12 presents the summary statistics of the sample, broken down by key demographic variables and risk factors for diarrhoea. Demographic and WASH characteristics appear to be similar between survey arms and stool collection groups within rounds. For example, in round one around 90% of respondents in all arms were female. In all arms close to 100% of respondents reported obtaining water from a tap. A majority in all arms in both rounds also reported storing water in a closed container. While the proportion of those who treated their water differed between rounds, they were similar between arms within rounds.

Table 12 Key Demographics

	Round One				Round Two			
	Standard (No Stool)	**Standard (Stool)**	**Pictorial (No Stool)**	**Pictorial (Stool)**	**Standard (No Stool)**	**Standard (Stool)**	**Pictorial (No Stool)**	**Pictorial (Stool)**
Total Respondents	147	25	145	31	163	35	140	31
% Female	91.8 [87.4, 96.3]	100.0 [86.3, 100.0]	91.03 [86.3, 95.7]	100.0 [88.8, 100.0]	77.3 [70.8, 83.8]	85.7 [73.5, 97.9]	75.0 [67.7, 82.3]	93.9 [70.2, 97.6]
Obtains Water from Tap	96.6 [93.6, 99.6]	96.0 [87.7, 100.0]	97.2 [94.5, 99.9]	100.0 [88.8, 100.0]	96.3 [93.4, 99.2]	100.0 [90.0, 100.0]	96.4 [93.3, 99.5]	96.8 [90.2, 100.0]
Stores Water in Closed Bucket	83.0 [76.8, 89.1]	84.0 [68.5, 99.4]	78.6 [71.9, 85.4]	93.5 [84.4, 100]	95.1 [91.7, 98.4]	91.4 [81.7, 100.0]	97.9 [95.4, 100.0]	93.5 [84.3, 100.0]
Does Something to Treat Water	25.2 [18.1, 32.3]	12.0 [0, 25.7]	21.4 [14.6, 28.1]	38.7 [20.5, 56.9]	68.1 [60.1, 75.3]	65.7 [49.2, 82.3]	71.4 [63.8, 79.0]	64.5 [46.7, 82.3]
Washes Hands with Water and Soap	75.5 [68.5, 82.5]	88.0 [74.3, 100.0]	79.3 [72.6, 86.0]	80.6 [65.9, 95.4]	96.9 [94.3, 99.6]	94.3 [86.2, 100.0]	97.1 [94.3, 99.9]	100.0 [88.8, 100.0]

Table 12 Caption: Percentages and 95% confidence intervals of key demographic and WASH indicators between rounds, arms, and stool collection status: *%[95%CI]*

Diarrhoea Rates for Each Proxy marker

Table 13 reports the diarrhoea rate estimated using data from each measure of diarrhoea. In round one, 37% [95%CI:30, 46] of participants in the standard survey reported diarrhoea, 61% [54, 68] of participants in the pictorial survey reported diarrhoea, 100% [94, 100] of stools had elevated levels of calprotectin (>7.25ug/g) or lactoferrin (>50ug/g) (or both), and 48% [34, 62] of stools were visually loose or watery. Similarly, estimated diarrhoea rates in the second round varied by measurement method, with the standard survey participants again reporting higher rates than those in the pictorial survey.

Table 13 Diarrhoea Rates Reported by Each Measure of Diarrhoea

	Round One	**Round Two**
Standard Survey	37 [30, 46]	26.9 [21, 34]
Pictorial Survey	61 [54, 68]	60 [52, 67]
Protein Marker	100 [94, 100]	57.8 [45, 70]
Visual Analysis	48 [34, 62]	51.5 [39, 64]

Table 13 Caption: Percentages and 95% Confidence Intervals of Each Measure of Diarrhoea Across Rounds: *%[95%CI]*

Infection Rates

Table 14 reports the infection rates of the different enteric pathogens – broken down by round, survey arm, and reported diarrhoea status. Overall, 64% [51, 77] of stools had at least one enteric pathogen detected in round one (wet period) and 79% [68, 89] had at least one enteric pathogen detected in round two (dry period). Bacterial pathogens were detected more often in round one (where 42% [28, 55] of stools had at least one bacterial pathogen) than round two (19% [9, 29]). On the contrary, viral pathogens were detected more often in round two (70% [58, 82]) than round one (38% [25, 51]). Enterotoxigenic E. Coli was the most common bacteria detected in both rounds, and Adenovirus was the most common virus detected in both rounds. Giardia was detected in 5% [0, 10] of stools in round one, and 8% [0, 15] of stools in round two.

Rates of enteric pathogen detection, both as a whole and for individual pathogens, were similar between diarrhoea statuses within arms and rounds.

Table 14 a and b Detection Rates of enteric pathogens detected in stool, broken down by round, survey arm, and reported stool status

Round One						
		Standard Survey		Pictorial Survey		
Type	**Species**	**Healthy (n=13)**	**Diarrhoea (n=12)**	**Healthy (n=14)**	**Diarrhoea (n=17)**	**Overall (n=56)**
	Enteropathogenic Escherichia Coli	23.1 [0.0, 49.5]	33.3 [2.0, 64.6]	0.0 [0.0, 3.3]	11.7 [0.0, 28.8]	15.4 [0.1, 0.3]
	Enterotoxigenic Escherichia Coli	23.1 [0.0, 49.5]	16.7 [0.0, 41.3]	35.7 [7.0, 64.4]	17.6 [0.0, 37.8]	22.8 [11.6, 34.0]
	Shiga Toxin Producing Escherichia Coli	0.0 [0.0, 3.3]	0.0 [0.0, 3.4]	0.0 [0.0, 3.3]	0.0 [0.0, 3.1]	0 [0.0, 2.7]
	Campylobacter	0.0 [0.0, 3.3]	0.0 [0.0, 3.4]	7.1 [0.0, 22.6]	11.7 [0.0, 28.8]	5.4 [0.0, 11.4]
	Salmonella	0.0 [0.0, 3.3]	16.7 [0.0, 41.4]	14.2 [0.0, 35.2]	0.0 [0.0, 3.1]	6.4 [0.0, 12.5]
	Shigella	0.0 [0.0, 3.3]	0.0 [0.0, 3.4]	7.1 [0.0, 22.6]	0.0 [0.0, 3.1]	1.66 [0.0, 4.9]
	Cholera	0.0 [0.0, 3.3]	0.0 [0.0, 3.4]	0.0 [0.0, 3.3]	0.0 [0.0, 3.1]	0 [0.0, 2.7]
Bacteria	**Any Bacteria**	**46.2 [14.7, 77.5]**	**50 [16.8, 83.1]**	**50.0 [20.0, 80.0]**	**29.4 [5.3, 53.6]**	**41.6 [28.4, 54.8]**
Protozoa	**Entamoeba histolytica**	0.0 [0.0, 3.3]	0.0 [0.0, 3.4]	0.0 [0.0, 3.3]	0.0 [0.0, 3.1]	0 [0.0, 2.7]

	Cryptosporidium	0.0 [0.0, 3.3]	0.0 [0.0, 3.4]	0.0 [0.0, 3.3]	0.0 [0.0, 3.1]	0 [0.0, 2.7]
	Giardia	0.0 [0.0, 3.3]	8.3 [0.0, 26.7]	14.2 [0.0, 35.3]	0.0 [0.0, 3.1]	4. 8 [0.0, 10.3]
	Any Protozoa	**0.0 [0.0, 3.3]**	**8.3 [0.0, 26.7]**	**14.2 [0.0, 35.3]**	**0.0 [0.0, 3.1]**	**4. 8 [0.0, 10.3]**
Viruses	**Rotavirus**	0.0 [0.0, 3.3]	0.0 [0.0, 3.4]	0.0 [0.0, 3.3]	0.0 [0.0, 3.1]	0 [0.0, 2.7]
	Sapovirus	0.0 [0.0, 3.3]	8.3 [0.0, 26.7]	7.1 [0.0, 22.6]	5.9 [0.0, 18.4]	5.4 [0.0, 11.4]
	Hepatitis E	0.0 [0.0, 3.3]	0.0 [0.0, 3.4]	0.0 [0.0, 3.3]	0.0 [0.0, 3.1]	0 [0.0, 2.7]
	Adenovirus	38.4 [7.9, 69.1]	16.7 [0.0, 41.4]	21.4 [0, 46.0]	11.8 [0.0, 29.8]	20.7 [10.0, 31.4]
	Astrovirus	0.0 [0.0, 3.3]	0.0 [0.0, 3.4]	0.0 [0.0, 3.3]	0.0 [0.0, 3.1]	0 [0.0, 2.7]
	Norovirus	15.3 [0.0, 38.1]	25 [0.0, 53.7]	28.6 [1.5, 55.6]	5.9 [0.0, 18.4]	16.7 [7.0, 26.4]
	Any Virus	**38.4 [7.9, 69.1]**	**41.6 [8.9, 74.4]**	**57.1 [27.5, 86.8]**	**23.5 [1.0, 46.0]**	**37.9 [25.0, 50.8]**
Any Infection		**69.2 [40.2, 98.3]**	**75.0 [46.3, 100]**	**78.5 [54.0, 100]**	**47.1 [20.6, 73.5]**	**64.2 [51.3, 77.3]**

Round Two *%[95%CI]*						
		Standard Survey		**Pictorial Survey**		
Type	**Species**	**Healthy (n=19)**	**Diarrhoea (n=13)**	**Healthy (n=18)**	**Diarrhoea (n=13)**	**Overall (n=63)**
	Enteropathogenic Escherichia Coli	0.0 [0.0, 3.0]	0.0 [0.0, 3.3]	16.7 [0, 35.7]	0.0 [0.0, 3.3]	5.0 [0.0, 10.1]
	Enterotoxigenic Escherichia Coli	15.8 [0.0, 33.8]	7.7 [0.0, 24.5]	11.1 [0, 27.1]	7.7 [0.0, 24.5]	11.6 [3.3, 19.8]
	Shiga Toxin Producing Escherichia Coli	0.0 [0.0, 3.0]	7.7 [0.0, 24.5]	5.6 [0.9, 17.3]	7.7 [0.0, 24.5]	4.5 [0.0, 9.7]
	Campylobacter	0.0 [0.0, 3.0]	15.3 [0, 38.1]	5.6 [0.0, 17.3]	7.7 [0.0, 24.5]	6.0 [0.0, 12.0]
	Salmonella	0.0 [0.0, 3.0]	7.7 [0.0, 24.5]	0 [0.0,3.1]	0.0 [0.0, 3.3]	1.3 [0.0, 4.0]
	Shigella	0.0 [0.0, 3.0]	0.0 [0.0, 3.3]	5.6 [0.0, 17.3]	7.7 [0.0, 24.5]	3.0 [0.0, 7.2]
	Cholera	0.0 [0.0, 3.0]	0.0 [0.0, 3.3]	0 [0.0,3.1]	7.7 [0.0, 24.5]	1.3 [0.0, 3.0]
Bacteria	**Any Bacteria**	**15.8 [0, 33.8]**	**15.3 [0, 38.1]**	**22.2 [0, 43.5]**	**23.1 [0, 49.6]**	**18.9 [9.0, 28.8]**
Protozoa	**Entamoeba histolytica**	0.0 [0.0, 3.0]	0.0 [0.0, 3.3]	0 [0.0,3.1]	0.0 [0.0, 3.3]	0.0 [0.0, 2.7]

	Cryptosporidium	0.0 [0.0, 3.0]	0.0 [0.0, 3.3]	0 [0.0,3.1]	0.0 [0.0, 3.3]	0.0 [0.0, 2.7]
	Giardia	5.3 [0.0, 16.3]	0.0 [0.0, 3.3]	11.1 [0.0, 27.2]	15.3 [0.0, 38.1]	7.8 [0.0, 14.6]
	Any Protozoa	**5.3 [0.0, 16.3]**	**0.0 [0.0, 3.3]**	**11.1 [0.0, 27.2]**	**15.3 [0.0, 38.1]**	**7.8 [0.0, 14.6]**
Viruses	**Rotavirus**	31.6 [8.5, 54.6]	7.7 [0.0, 24.5]	33.3 [9.2, 57.5]	7.7 [0.0, 24.5]	23.8 [12.8, 34.7]
	Sapovirus	0.0 [0.0, 3.0]	0.0 [0.0, 3.3]	0 [0.0,3.1]	0.0 [0.0, 3.3]	0.0 [0.0, 2.7]
	Hepatitis E	0.0 [0.0, 3.0]	0.0 [0.0, 3.3]	0 [0.0,3.1]	0.0 [0.0, 3.3]	0.0 [0.0, 2.7]
	Adenovirus	42.1 [17.7, 66.6]	53.8 [33.5, 85.2]	44.4 [19.0, 69.9]	61.5 [30.9, 92.1]	48.2 [35.6, 60.9]
	Astrovirus	0.0 [0.0, 3.0]	0.0 [0.0, 3.3]	0 [0.0,3.1]	0.0 [0.0, 3.3]	0.0 [0.0, 2.7]
	Norovirus	31.6 [8.5, 54.6]	30.7 [1.7, 59.8]	22.2 [0, 43.5]	7.7 [0.0, 24.5]	24.5 [13.5, 35.5]
	Any Virus	**68.4 [45.4, 91.4]**	**69.2 [40.2, 98.3]**	**72.2 [49.3, 95.1]**	**69.2 [40.2, 98.3]**	**69.83 [58.2, 81.5]**
Any Infection		**73.7 [51.9, 95.5]**	**76.9 [50.4, 100]**	**77.8 [56.5, 99.1]**	**92.3 [75.6, 100]**	**78.7 [68.3, 89.2]**

Table 14 a and b Caption: Detection rates of enteric pathogens detected in stool samples collected by PCR, broken down by round, study arm, and diarrhoea status: *%{95%CI]*

Diagnostic Performance Indicators

Table 15 and figure 16 present the sensitivities and specificities of each measure of diarrhoea as a proxy marker of enteric pathogens being detected in stool. When looking at the presence of at least one enteric pathogen in the stool, proteins had the highest sensitivity, 0.73 [0.61, 0.82]. This indicates that 73% of those with at least one enteric pathogens detected in their stool had a positive test (in this case, elevated protein levels). This compares to sensitivities of 0.49 [0.32, 0.66] for the standard survey, 0.40 [0.25, 0.57] for the pictorial survey, and 0.46 [0.34, 0.57] for visual confirmation. There was little evidence of differences in sensitivity across the four proxy markers when broken down by pathogen type.

Regarding specificity of the proxy markers as markers of at least one enteric pathogen being detected in stool, proteins had the lowest, 0.18 [0.07, 0.31]. This indicates that 18% of those without at least one enteric pathogen in their stool had a negative test (in this case, normal protein levels). This compares to specificities of 0.65 [0.41, 0.85] for the standard survey, 0.36 [0.17, 0.59] for the pictorial survey, and 0.43 [0.27, 0.59] for visual confirmation. Again, There was little evidence of differences in specificities when breaking down by pathogen type. Overall, I also see a trade-off between sensitivity and specificity. This is represented by the 45 degree line in figure 16.

Table 15 Performance of Diarrhoea Measurement Methods as Indicators of Enteric Infection

		Standard Survey	Pictorial Survey	Visual Confirmation	Proteins
All	**Sensitivity**	0.49 [0.32, 0.66]	0.40 [0.25, 0.57]	0.46 [0.34, 0.57]	0.73 [0.61, 0.82]
	Specificity	0.65 [0.41, 0.85]	0.36 [0.17, 0.59]	0.43 [0.27, 0.59]	0.18 [0.07, 0.31]
Bacteria	**Sensitivity**	0.48 [0.23, 0.72]	0.42 [0.20, 0.66]	0.56 [0.38, 0.73]	0.84 [0.68, 0.94]
	Specificity	0.56 [0.41, 0.73]	0.49 [0.33, 0.65]	0.53 [0.42, 0.64]	0.28 [0.18, 0.38]
Virus	**Sensitivity**	0.45 [0.27, 0.64]	0.44 [0.27, 0.62]	0.51 [0.38, 0.63]	0.72 [0.60, 0.83]
	Specificity	0.56 [0.35, 0.76]	0.45 [0.33, 0.58]	0.51 [0.37, 0.65]	0.18 [0.08, 0.30]
Protozoa	**Sensitivity**	0.50 [0.01, 0.99]	0.33 [0.04, 0.78]	0.50 [0.16, 0.84]	0.75 [0.35, 0.97]
	Specificity	0.56 [0.41, 0.69]	0.46 [0.32, 0.59]	0.50 [0.40, 0.60]	0.23 [0.15, 0.32]

Table 15 Caption: Diagnostic Performance Indicators [with 95%Cis] of each diarrhoea measure as a proxy marker of enteric infection

Figure 16 ROC Space Plot of Diarrhoea Measurement Methods as Markers of Enteric Infection

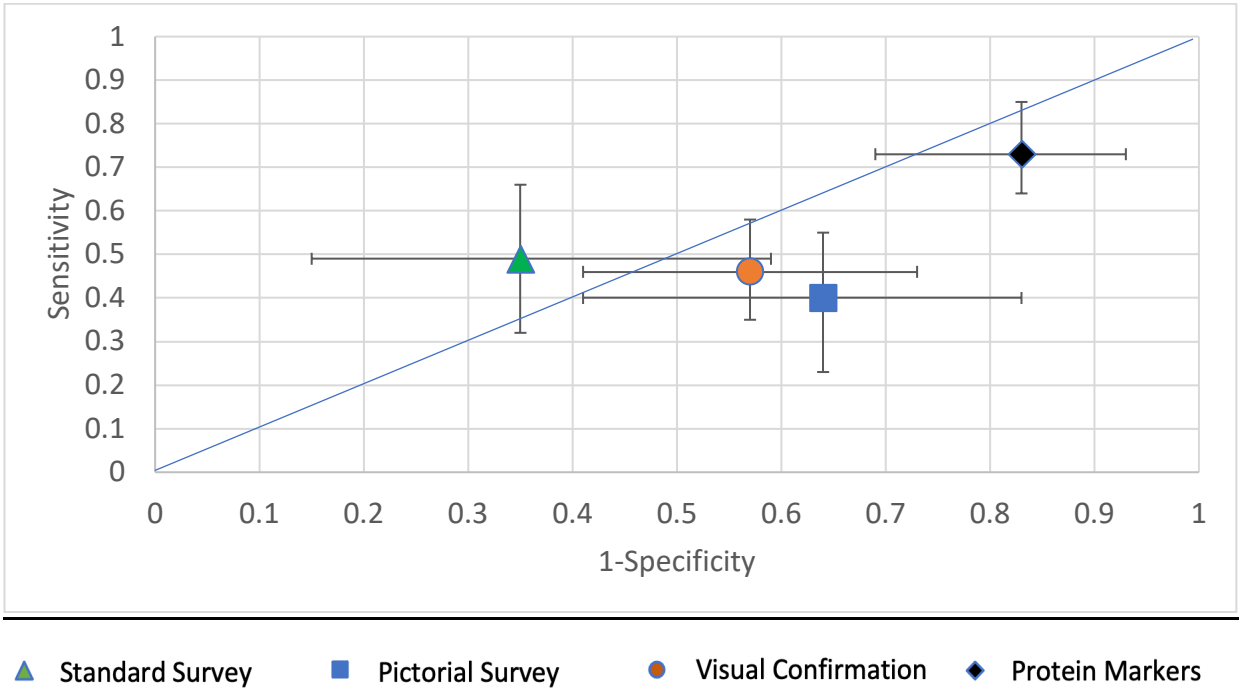


Figure 16 Caption: An ROC Space Plot showing the sensitivities and specificities of all four diarrhoea measurement methods as proxy markers of at least one enteric pathogen being present in stool. The 45 degree line represents the trade off between sensitivity and specificity, where along the line the proxy marker gives a 50% chance of obtaining the correct classification, with points towards the top left giving a better chance.

Measures of Agreement

Table 16 presents the agreement between the different proxy markers as measures of diarrhoea. No pairwise combination of the four proxy markers showed good agreement. For all pairwise combinations, the expected agreements (agreement expected by chance alone) were similar to the observed agreements[98]. This is supported by my estimates of kappa values – the relationship between observed and expected agreements. Negative or close to zero kappa values indicate no agreement, with values between 0 and 0.20 indicating only slight agreement [98]. All of my kappa value estimates are under 0.20, with most being negative.

Table 16 Measures of Agreement of the Different Proxy markers

Pairwise Combination	**Kappa (p value)**	**Observed Agreement (%)**	**Expected Agreement (%)**
Visual v Standard Survey	0.14 (0.14)	55.8	48.7
Protein v Standard Survey	0.02 (0.43)	48.2	47.3
Visual v Pictorial Survey	-0.12 (0.83)	44.3	50.2
Protein v Pictorial Survey	-0.05 (0.67)	46.8	69.1
Visual v Protein	-0.10 (0.89)	45.2	50.2

4.5.Discussion

My results indicate that in Cox's Bazar, diarrhoea, measured by any of my four different methods (standard, pictorial, visual, and protein), misclassifies diarrhoea when used as a proxy marker. This is also likely the case in environments similar to Cox's Bazar, in terms of both the epidemiology of enteric pathogens and possibly human behaviour. In fact, given that for each proxy marker Sensitivity+Specificity is approximately equal to 1, measuring diarrhoea by any of these methods is no better a proxy marker of enteric infection than flipping a coin (if this was higher than one, it would be a better proxy marker than flipping a coin, and if lower than 1 a worse proxy marker)[81]. I further found that there was little to no agreement between the four methods of diarrhoea measurement. This indicates that estimates of diarrhoea differ by the methods used – supporting findings in chapters Two and Three.

Impact on WASH Trials

Most WASH trials to date have used carer-reported diarrhoea as their main endpoint. While it is important to acknowledge that WASH interventions aim to reduce the burden of diarrhoea, they do not aim to reduce the burden of all types of diarrhoea. By forming a barrier between pathogens in the environment and susceptible people, WASH interventions work to reduce the transmission of diarrhoea causing enteric pathogens, and thus pathogenic-diarrhoea. Diarrhoea has causes other than enteric-infection though, such as nutritional intolerances, adverse drug reactions, non-enteric infections, and chronic bowl problems such as tropical sprue[40, 99-101]. WASH interventions cannot be reasonably expected to reduce the burden of diarrhoea caused by these (with perhaps the exception of tropical sprue in the long run). Further, there should be an interest in asymptomatic enteric infection. Even though asymptomatic enteric infections may not expressly cause disease and death, it has still been associated with stunting[39]. Further, those with an asymptomatic infection may pass on infection to somebody who may

experience severe morbidity or mortality. The use of diarrhoea as the main endpoint rather than infection precludes the measurement of these cases.

Misclassification error when using diarrhoea as a proxy marker of enteric infections in WASH trials may explain the null findings found in many WASH trials, rather than true intervention ineffectiveness. Two of the three recent large scale WASH trials (the two WASH Benefits and one SHINE trial), which used carer reported diarrhoea with a seven day recall as their endpoint, reported null results[78-80]. Misclassification error resulting in null intervention effects, despite there being a true intervention effect, can be illustrated quantitatively. I will look at Null et al.'s (2018) WASH Benefits trial in rural Bangladesh[80]. The WASH Benefits trial in Bangladesh was one of the largest integrated WASH trials in history, enrolling 8246 households in 702 clusters[80]. The WASH Benefits trial in Bangladesh concluded that the combined intervention did not impact carer-reported diarrhoea rates (Adjusted Prevalence Difference=0.7% [95%CI:-2.4, 3.7]). Greenland (1996) showed that misclassification error in the measurement of dichotomous outcome variables (such as the misclassification of infection by using diarrhoea), particularly when this misclassification error results in poor specificity, biases estimated intervention effects towards the null even when misclassification is equal in the control and intervention arms[81]. I demonstrate graphically the impact that varying sensitivities and specificities can have on reported infection relative risks when using a proxy marker, such as diarrhoea (Fig. 17). If the intervention had a true relative risk of enteric infection of 0.5, the intervention halving the risk of enteric infection, but used a proxy marker with sensitivities and specificities of around 50% (as I observed in Cox's Bazar), the relative risk estimated would be around 1 – a null intervention effect, as was seen in the trial.

Figure 17 The Impact of Sensitivity and Specificity on Intervention Effects

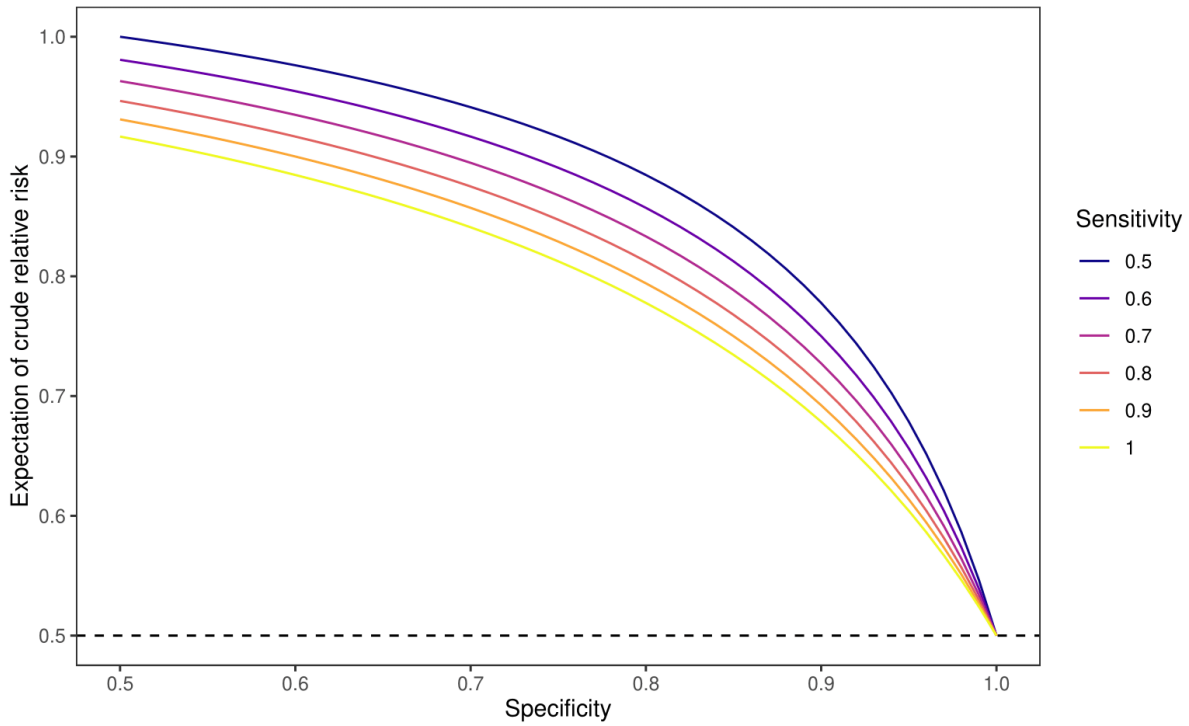


Figure 17 Caption: Relationship between outcome classification error and the expected value of the crude relative risk estimator. The prevalence of infection in the infection and control groups are 20% and 10%, respectively, so the true relative risk is 0.5.

Impact on Surveillance Activities

In addition to possibly causing bias in the results of WASH trials, diarrhoea being a poor proxy marker of infection may bias results from surveillance programs. These programs aim to understand where infection lies to deploy interventions (such as water or hygiene systems) and detect outbreaks of enteric infection. However, the impacts of sensitivity and specificity vary depending on underlying infection prevalence. In high prevalence environments, such as Cox's Bazar, sensitivity has a more dramatic impact on reported rates than specificity (Fig. 18). Generally, misclassification in high prevalence environments shifts estimates of prevalence down[102]. As most people have an infection, there are less non-infectious cases to rule out. However, there are more infectious cases to include. In an area like Cox's Bazar with high levels of infection prevalence the sensitivity of proxy markers should be prioritised. This means that the measurement of proteins as a proxy marker of infection may be suitable (at least for infection surveillance, not WASH trials). However, in low prevalence environments, such as London, specificity is far more important, with misclassification reporting in inflated estimates (Fig. 18)[102]. In low prevalence environments specific tests should be prioritised.

Figure 18 The Impact of Sensitivity and Specificity on Reported Disease Rates

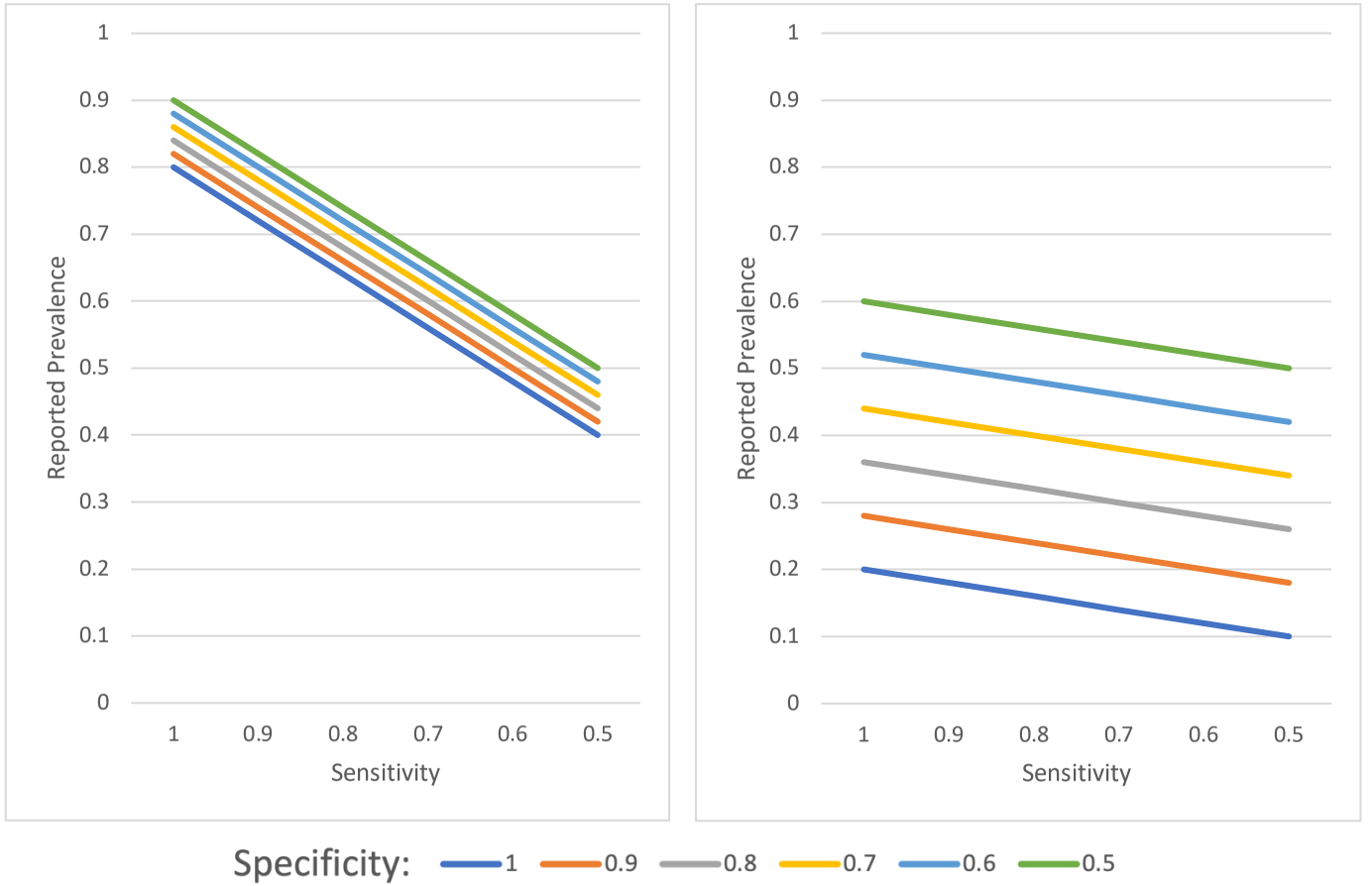


Figure 18 Caption: Relationship between outcome classification error and the expected prevalence: a (left): A true prevalence of 80%; b (right): A true prevalence of 20%

Sources of Misclassification

Non-Pathogenic Diarrhoea Resulting in Decreased Specificity

There are multiple reasons why a child without an enteric infection may have diarrhoea. Nutritional intolerances can cause diarrhoea, with common intolerances including those to certain sugars and proteins[99]. Certain clinical treatments and medications can also increase the risk of diarrhoea, with up to a quarter of under-fives on some antibiotics having diarrhoea[100]. Health conditions that are not expressly to do with the gut can also be associated with diarrhoea, such as HIV, in which around a third of those afflicted have non-infectious diarrhoea due to associated enteropathy[40]. Finally, chronic bowel problems, such as tropical sprue, can result in diarrhoea regardless of infection[101]. However, I do not think that these factors explain more than a small proportion of the lack of agreement observed here, as most children in Cox's Bazar do have some form of enteric infection. In a low prevalence environment, however, these factors would be of more importance. As they impact specificity, they are also important in WASH trials.

Asymptomatic Infections Resulting in Decreased Sensitivity

My results support past work that has found high degrees of asymptomatic infection. Okitsu et al. (2020) found that 80% of healthy children in rural Bangladesh have some enteric infection[38]. Oberhelman et al. (2001) similarly reported that rates of Salmonella infection were higher than carer-reported diarrhoea rates, indicating a degree of asymptomatic infection[103]. Prado et al. (2005) also obtained similar results for Giardia[104]. I suspect that asymptomatic infection is a major source of misclassification in high prevalence environments such as Cox's Bazar.

Poor reliability

With respect to the three visual inspection based methods (standard question, stool chart and third party inspection) I think another mechanism is in play – the unreliability of human adjudication. The clue to this explanation lies in the literature

and in the data from this study. I have demonstrated in this study that there is little to no agreement between the different methods. In my systematic review, I also found that diarrhoea measurement methods are highly sensitive to their methods, and may even be reactive to being observed or taking part in an interventional trial (Chapter Two). This undoubtably also results in misclassification – at least for the standard and pictorial methods where carers are questioned.

Limitations

My study has several limitations. First, despite my extensive panel, I may have omitted some endemic diarrhoea causing pathogens (perhaps some yet to be discovered). If survey diarrhoea rates were more sensitive predictors of these omitted organisms than of those included, then this would somewhat improve the low sensitivities and specificities observed here. Second, I did not quantify the amount of pathogen present. WASH interventions may reduce the load of infectious agents in an infected child as well as the proportion of children who are infected at all. Third, I did not analyze the prevalence of co-infection which might also be reduced by WASH interventions. Fourth, the number of stool analyses is small, though my results suggest low test accuracy even at the limits of the 95% confidence interval. Fifth, my study was done in one center only and hence should be replicated (though we find it hard to see why the results should not apply in similar settings). Sixth, I deliberately manipulated the sampling so that we would have equal numbers of samples from participants with and without reported diarrhoea, thereby biasing prevalence and vitiating direct measurement of predictive values. That said, my estimates of sensitivity and specificity could be applied to populations of known prevalence under the assumption of independence of these parameter types. Against these limitations our study has strengths in terms of multiple measures of diarrhoea rates and use of ‘gold standard’ PCR testing following internationally recognized standards.

Conclusion

I have empirically demonstrated that diarrhoea, measured through any means, is not a good proxy marker of enteric infection, at least in Cox's Bazar and other similar environments. It is likely that this problem is more acute and impactful in environments with high a prevalence of enteric infection, and less so (but still important) in environments with a lower prevalence. It is also important to stress that my findings apply to active surveillance and not to cases that present clinically in health facilities. Clinical cases make up only 2% of cases detected by active surveillance[27, 28]. This means that my findings are silent on clinical cases. It is often stated that sensitivity and specificity, unlike predictive values, are not affected by prevalence. But this only holds true when sensitivity does not change in line with prevalence. In the case of diarrhoea there is almost certainly a correlation between prevalence and sensitivity – at the limit, people with dysentery or rice water stools will almost certainly have a micro-biological cause; hence my strong caution against extrapolating my findings to the clinic.

Traditional means of mass stool testing for active surveillance, such as polymerase chain reaction (PCR), can be cost and resource prohibitive in many settings. However, there are several alternative methods. Samples can be pooled to cut down on the number of tests needed, particularly for low-prevalence infections. Recent work has also shown the utility of testing of faecal sludge instead of stool samples, in essence a type of sample pooling which also cuts down on sample collection difficulties. Further, the use of rapid diagnostic tests (RDTs) can be used in lieu of PCR, cutting down on the cost of shipment, processing, and personnel; and allowing for use in low resource settings. It is important to note however that RDTs, and other tests that are not the 'PCR gold standard' also suffer from sensitivity and sensitivity issues as measures of infection (albeit at a much lower level than diarrhoea). The WHO recommends that RDTs for Cholera, for example, should have a minimum sensitivity of 90% and a minimum sensitivity of 85%[15]. This may still result in biased intervention effects and prevalence (especially in low

prevalence areas), however unbiased estimates can be obtained with appropriate statistical modelling, which is much easier when the sensitivity and specificity are known. In the next chapter, I discuss how results from stool testing can be used in public health programming.

5. Uses of Enteric Infection Surveillance: Prevalence and the Spatio-Temporal Distribution of Common Enteric Pathogens in the Cox's Bazar Forcibly Displaced Myanmar Nationals Camp: A Cross-Sectional Observational Study

5.1.Abstract

Introduction

Diarrhoea, commonly caused by enteric infections, is a significant cause of under-five morbidity and mortality in refugee camps. Knowledge on which enteric pathogens are most prevalent and their spatio-temporal distribution is needed deploy interventions to combat their spread. Using data collected from Chapter Four, explored these in a portion of the Cox's Bazar FDMN Camp Complex through a retrospective analysis.

Methods

I collected 120 stool samples, split evenly between the wet and dry season, in the Leda Makeshift Camp area of Cox's Bazar, an approximately 1km^2 area with a population of around 14,000 as described in Chapter Four. Stool samples were tested for 16 enteric pathogens and two proteins associated with enteric infection. I estimated the prevalence of common pathogens in each season, and analysed their spatio-temporal distribution. I also analysed the relationship between household distance from latrines and water taps and infection, as this may be an important factor in their spatial distribution.

Results

The most common pathogens found were Enterotoxigenic Escheria Coli (23% [95%CI:12, 34] point prevalence in the wet season, 12% [3, 20] in the dry); Enteropathogenic Escheria Coli (16% [0, 0] in the wet, 5% [0, 10] in the dry); Rotavirus (0% in the wet, 23% [13, 35] in the dry); Adenovirus (21% [10, 31] in the wet, 48% [36, 61] in the dry); and Norovirus (16% [7, 26] in the wet, 25% [14, 36] in the dry). Viral infections were 8 [23, 7] percentage points (pp) less common in the wet season than dry, and bacterial infections were 23 [17, 39] pp more common in the wet season than dry. I found no evidence that infections were clustered spatially or temporally within

the study site. There was additionally no evidence of a relationship between household distance from water sources and latrines and most enteric pathogens.

Implications

The pathogens found in the FDMN community living in the Leda Makeshift Camp area of Cox's Bazar are similar to those found previously in the wider Bangladeshi community. On the small scale of my study area, there is little evidence in heterogeneity in infection risk – suggesting common exposure to sources of infection. Infection control policy and program should be consistent in these small areas, and similar in both refugee and host populations. They should be altered by season, however.

5.2.Introduction

Diarrhoea accounted for 7% of all under-five deaths in United Nations High Commissioner for Refugees (UNHCR) registered refugee camps between 2006 and 2010[105]. Children in refugee camps are at increased risk of diarrhoea given the ease at which diarrhoea causing enteric pathogens spread in these environments, due to dense populations, insufficient water and sanitation infrastructure, and poor hygiene[106]. To combat the spread of enteric pathogens, camp authorities often implement WASH interventions, as well as clinical interventions such as vaccination, nutritional supplementation, and treatment once symptomatic.

As discussed in Chapter Four, it is important to look at infection, not disease, when evaluating trials and performing disease surveillance. Additionally, as different pathogens have different control techniques, knowledge on which enteric pathogens are endemic in refugee camps allows for targeted, pathogen specific interventions to prevent spread. For example, the control of soil-transmitted helminths relies primarily on deworming programs and the prevention of geophagy[107]. These interventions, however, would likely have little impact on cholera, where vaccination programs and the provision of safe water are the best solutions[108]. Knowledge on which enteric pathogens are present, as well as their epidemiology, generally does not exist in refugee camps – only in host communities. While it is possible that host communities may have similar enteric pathogen profiles to refugee camps, that is unclear and has not been researched. The global enteric multicentre (GEMS) study, for example, identified Rotavirus, Cryptosporidium, Enterotoxigenic Escheria. Coli, Shigella, Adenovirus, Salmonella, and Campylobacter as some of the most prevalent enteric infections in parts of Bangladesh, but did not include refugees in their sample, despite Bangladesh being home to over 1 million refugees[48].

Similarly, no studies have examined the spatio-temporal distribution of enteric pathogens in refugee camps – with studies to date only measuring the spatio-temporal distribution of diarrhoea[105]. In addition to diarrhoea being a poor marker of enteric infection, as described in Chapter Four, looking at diarrhoea alone precludes knowledge on which pathogens are causing the diarrhoea. While other studies have made inferences on the spatio-temporal distribution of enteric pathogens outside of refugee camps – including an increased risk of bacterial infection in the rainy season, and an increased risk of viral infection in the dry season, it is unclear how generalisable these findings are, particularly to a refugee camp setting, due to differences in the population, environment, and overall socioeconomic context[109]. However, knowing the epidemiology of enteric pathogens in refugee camps, particularly the pathogens present and their spatio-temporal distribution, enables program and policymakers to be able to better design and deploy interventions to control pathogens.

Using data collected in Chapter Four, I performed an exploratory retrospective analysis to examine the prevalence and spatio-temporal distribution of 16 common enteric pathogens among children under-five years in the Leda Makeshift Camp area of the Cox's Bazar Forcibly Displaced Myanmar Nationals (FDMN) Camp in Bangladesh. I also examined the same for calprotectin and lactoferrin, two proteins associated with enteric infections. By examining for calprotectin and lactoferrin, we can see not only where infection currently lies, but may have lied in the recent past. Calprotectin and lactoferrin are associated with recent bacterial infection, as well as persistent infections which result in environmental enteropathy[50, 51]. Understanding the limitations introduced by using pre-existing data, the objectives were to document: (1) the causes of enteric infection in the camp; (2) levels of calprotectin and lactoferrin (proteins that are associated with enteric infections)[50, 51]; and (3) the temporal and spatial distribution of enteric pathogens and elevated protein levels in the camp (with

particular regard to household distance from latrines and water taps, and use of WASH interventions).

5.3.Methods

Study Design

The design and data collection for this study has been previously described in Chapter Four Section 3 (Methods). I aimed to recruit 800 participants, split equally between two rounds: round one in the wet season (July/August 2019), and round two in the dry season (November/December 2019). I randomised the participants equally into two arms of data collection, with each arm differing by how questions about were asked to carers. Stool was collected from random subset, split equally between data collection arm, diarrhoea status, and season. Stools were tested for 16 endemic pathogens (see Chapter Four for methods and a list of pathogens). I also tested stools for two proteins associated with bacterial enteric infection (lactoferrin and calprotectin), as this captures pathogens which I may not have tested for, particularly bacteria[50, 51]. For this study, I included only the observations from which I had collected stool samples.

Data Analysis

I described the means and 95% confidence intervals of key demographic variables and indicators of WASH system usage, as well as the prevalence of pathogens and elevated protein levels (defined as a faecal calprotectin level over 50ug/g and/or a faecal lactoferrin level over 7.25ug/g), broken down by rainy or dry season[50, 51]. I also estimated the differences in prevalence between seasons for each pathogen and elevated protein levels (with 95% confidence intervals). To determine if there was clustering by weeks within seasons, I graphically plotted rates of each pathogen and elevated protein levels against week, with 95% confidence intervals. Using an ANOVA analysis, I also compared differences in these rates between weeks with the null hypothesis of no difference between weeks within rounds against the alternative that at least one week had a different prevalence.

I plotted on a map of the camp the geolocation and infection status of each stool for the most common enteric pathogens, as well as if the stool had elevated calprotectin levels, by survey week. To examine/explore clustering within seasons I calculated Moran's I statistic, a measure of spatial correlation which determines how similar neighbouring observations are compared to non-neighbouring observations[110]. Considering that the furthest point between any two observations was 0.5KM, I defined neighbouring observations as observations that were within as 0.1KM of each other, otherwise known as the threshold distance. This ensured that every observation had at least five neighbours, without every other observation being a neighbour. I also plotted the relationship between household distance from latrines and water taps and infection using a fractional polynomial plot with 95% confidence intervals.

All analyses were conducted with StataSE Version 15 and R 4.0.2. As stool samples were collected at a 1:1 ratio with regards to diarrhoea status, all sample statistics were weighted back to the proportions of cases with and without diarrhoea in the larger random sample. As I made use of pre-existing data without many variables of interest and a low sample size, I was unable to perform adjusted multivariable analyses.

Ethics

Ethical permission was granted by the University of Warwick Biomedical Science Research Ethics Committee (REGO-2019-2345) and by the ICDDR,B Ethical Review Committee (PR-19027). The study was also approved by the Refugee Relief and Repatriation Commission, Government of Bangladesh (#789) and Camp 24 authorities.

Full written consent was obtained, with a full data management and data collection plan in place. No identifiable data were shared with government authorities, UN agencies, or NGOs. No compensation was offered. The study was registered on ISCRTN during its inception, and updated as necessary (ISRCTN41564300).

5.4.Results

Study Recruitment

The full recruitment process is described in Chapter Four. In round one (wet season), I approached 68 households to provide stool samples, of which 56 provided samples. In round two (dry season), 102 households were approached and 63 provided stool samples. These households (n=119) form the basis for my analysis, and their demographic characteristics are described in Table 17.

Table 17 Demographic and WASH Usage Indicators

	Wet Season (Round 1)	**Dry Season (Round 2)**
Total Stool Samples Collected *n*	56	63
Female Respondents *n(%[95%CI])*	56 (100 [94, 100])	56 (85 [74, 92])
Mean Respondent Time in Bangladesh *months[95%CI]*	161 [141, 180]	206 [186, 225]
Mean Child Age *months[95%CI]*	30 [26, 35]	31 [27, 35]
Drinking Water Source		
Piped water to tap	55 (98 [89, 100])	65 (98 [89, 100])
Borehole	0 (0 [0, 3])	1(2 [0,11])
Dug well	0 (0 [0, 3])	0 (0 [0, 3])
Spring	0 (0 [0, 3])	0 (0 [0, 3])
Rain	0 (0 [0, 3])	0 (0 [0, 3])
Cart	1 (2 [0, 11])	0 (0 [0, 3])
Mean Water Wait Time *minutes[95%CI]*	83 [71, 96]	35 [27,4 3]
Does something to make water safe *n(%[95%CI])*	15(27 [17, 40])	43 (65 [53, 76])
Toilet unsafe day *n(%[95%CI])*	49(88 [76, 94])	65 (98 [90, 100])
Toilet unsafe dark *n(%[95%CI])*	29(52 [39, 65])	62 (94 [85, 98])
How Wash Hands *n(%[95%CI])*		
Does not wash hands	0 (0 [0, 3])	0(0)
Water alone	9(16 [9,28])	2 (3 [0, 11])
Water and soap	47(84 [72, 91])	64 (97 [89, 100])
Alcohol based hand sanitizer	0 (0 [0, 3])	0 (0)
Handwashing at all critical times *n(%[95%CI])*	53(96 [85, 98])	51 (77 [66, 86])
Water Storage *n(%[95%CI])*		
Open bucket	4 (7 [3, 18])	1 (2 [0, 10])
Closed bucket	50 (89 [78, 95])	61 (92 [83, 97])
Bottle	0 (0 [0, 3])	0 (0 [0, 3])
Jerry can	2 (4 [1, 13])	4 (6 [2, 15])
Breastfeeding *n(%[95%CI])*		
No	2 (4 [1, 14])	10 (15 [8, 26])
Yes - Partially	10 (18 [10, 30])	9 (14 [7, 24])
Yes - exclusively	44 (78 [66, 87])	47 (71 [59, 81])

Spatio-temporal distribution of Enteric Infection

Temporal Distribution

Table 18 presents the pathogens found in each season. In the wet season, 64.2% [95%CI:51.3, 77.3] of stools collected had at least one enteric pathogen detected. This compares to 78.7% [68.3, 89.2] in the dry season.

Detection of bacterial pathogens in stools differed significantly by season: 22.7 [16.5, 39.2] percentage points (pp) more stools had at least one bacterial pathogen detected in the wet season than the dry season. The most common bacterial enteric pathogen detected in both seasons was Enterotoxigenic Escherchia Coli (ETEC), with a 22.8% [11.6, 34.0] estimated prevalence in the wet season and a 11.6% [3.3, 19.8] estimated prevalence in the dry season. Salmonella, Campylobacter, and Enterpathogenic Escherchia Coli (EPEC) also varied between seasons in the same manner. Shiga toxin producing Escherchia Coli (STEC) was the only bacteria estimated to be more prevalent in the dry period than the wet (4.5pp [0, 9.7]). With the exception of EPEC, none of the bacteria had significant variations between weeks within seasons. In the wet season, 60% of stools had EPEC in week 29 compared to none in week 33 (p=0.03) (Fig. 19).

Detection of viral pathogens also differed significantly by season: 32.0pp [-49.4,-14.5] fewer stools had at least one viral infection in the wet season than the dry. The most common virus detected in stools in both seasons was adenovirus, which also had a lower estimated prevalence in the wet season (20.7% [9.96,31.36]) than the dry season (48.2% [35.61,60.87]). Adenovirus was detected in all weeks, with little difference in estimated prevalence between weeks within seasons (Fig. 19). Norovirus was also widespread, with an estimated prevalence of 16.7% [7.0,26.4] in the wet season and 24.5% [13.5,35.5] in the dry season. As with adenovirus, there was evidence of seasonality, but no evidence of differences in estimated prevalence between weeks within seasons (Fig. 19). Rotavirus was only detected in stools collected in the dry

season (23.8% [12.8,34.7]), and Sapovirus was only detected in stools collected in the wet season (5.4% [0.00,11.4]).

The only parasite detected was Giardia, which did not differ significantly by season. I found Giardia at an estimated prevalence of 4.8% [0.0, 10.3] in the wet season and 7.8% [0.0, 14.6] in the dry season. The prevalence of Giardia did not differ significantly by week within seasons (Fig. 19).

Elevated calprotectin levels were detected in all stools collected in the wet season (100% [93.6, 100]), and in 48.9% [35.9,61.4] of the stools collected in the dry season. Calprotectin levels appeared to decrease week by week in the dry season, but with some uncertainty (Fig. 19).

Table 18 Enteric Pathogens by Season

Type	Species	Round One (Wet Season) *%[95%CI]*	Round Two (Dry Season) *%[95%CI]*	Diff. Between Seasons *ppd[95%CI]*
Bacteria	Shiga toxin-producing Escherichia coli	0 [0, 2.7]	4.5 [0,9.7]	-4.5 [0,-9.7]
	Enteropathogenic Escherichia coli	15.4 [0.1,0.3]	5.0 [0.0,10.1]	10.5 [-0.6,0.21.5]
	Enterotoxigenic Escherichia Coli	22.8 [11.6,34.0]	11.6 [3.3,19.8]	11.4 [-2.7,25.2]
	Salmonella	6.4 [0.0,12.5]	1.3 [0.0,4.0]	5.0 [-1.7,11.8]
	Campylobacter	5.4 [0.0, 11.4]	6.1 [0.1,12.0]	0.7 [0.0,9.1]
	Shigella	1.7 [0.0,4.9]	3.0 [0.00,7.2]	-1.3 [-6.7,4.0]
	Cholera	0 [0, 2.7]	1.3 [0.0,3.0]	-1.3 [-4.0,1.3]
	TOTAL BACTERIA	**41.6 [28.4,54.8]**	**18.9 [9.0,28.8]**	**22.7 [16.5,39.2]**
Protozoa	Entamoeba histolytica	0 [0, 2.7]	0 [0, 2.7]	0 [0, 2.7]
	Cryptosporidium	0 [0, 2.7]	0 [0, 2.7]	0 [0, 2.7]
	Giardia	4. 8 [0.00,10.3]	7.8 [0.01,14.6]	-3.0 [-11.7,5.7]
	Any PROTOZOA	**4.8 [0.00,10.3]**	**7.8 [0.01,14.6]**	**-3.0 [-11.7,5.7]**
Viruses	Rotavirus	0 [0, 2.7]	23.8 [12.8,34.7]	-23.8 [-34.7,-12.8]
	Sapovirus	5.4 [0.00,11.4]	0 [0, 2.7]	5.4 [0.00,11.4]
	Hepatitis E	0 [0, 2.7]	0 [0, 2.7]	0 [0, 2.7]
	Adenovirus	20.7 [10.0,31.4]	48.2 [35.6,60.9]	-27.6 [-44.2,-11.0]
	Astrovirus	0 [0, 2.7]	0 [0, 2.7]	0 [0, 2.7]
	Norovirus	16.7 [7.0,26.4]	24.5 [13.5,35.5]	-7.8 [-22.5,6.9]
	Any Virus	**37.9 [25.0,50.8]**	**69.83 [58.2,81.5]**	**-32.0 [-49.4,-14.5]**
ANY INFECTION		**64.2 [51.3,77.3]**	**78.7 [68.3,89.2]**	**-14.4 [-31.2,2.3]**
Elevated Calprotectin		100 [93.6, 100]	48.9 [35.9,61.4]	51.3 [38.5,64.1]
Elevated Lactoferrin		51.9 [38.2,65.56]	32.57 [20.6,44.5]	19.3 [1.1,37.6]
Any Elevated Protein		**100 [93.6, 100]**	**57.17 [44.6,69.8]**	**42.8 [30.2,55.5]**

Figure 19 Weekly Prevalence of Common Pathogens

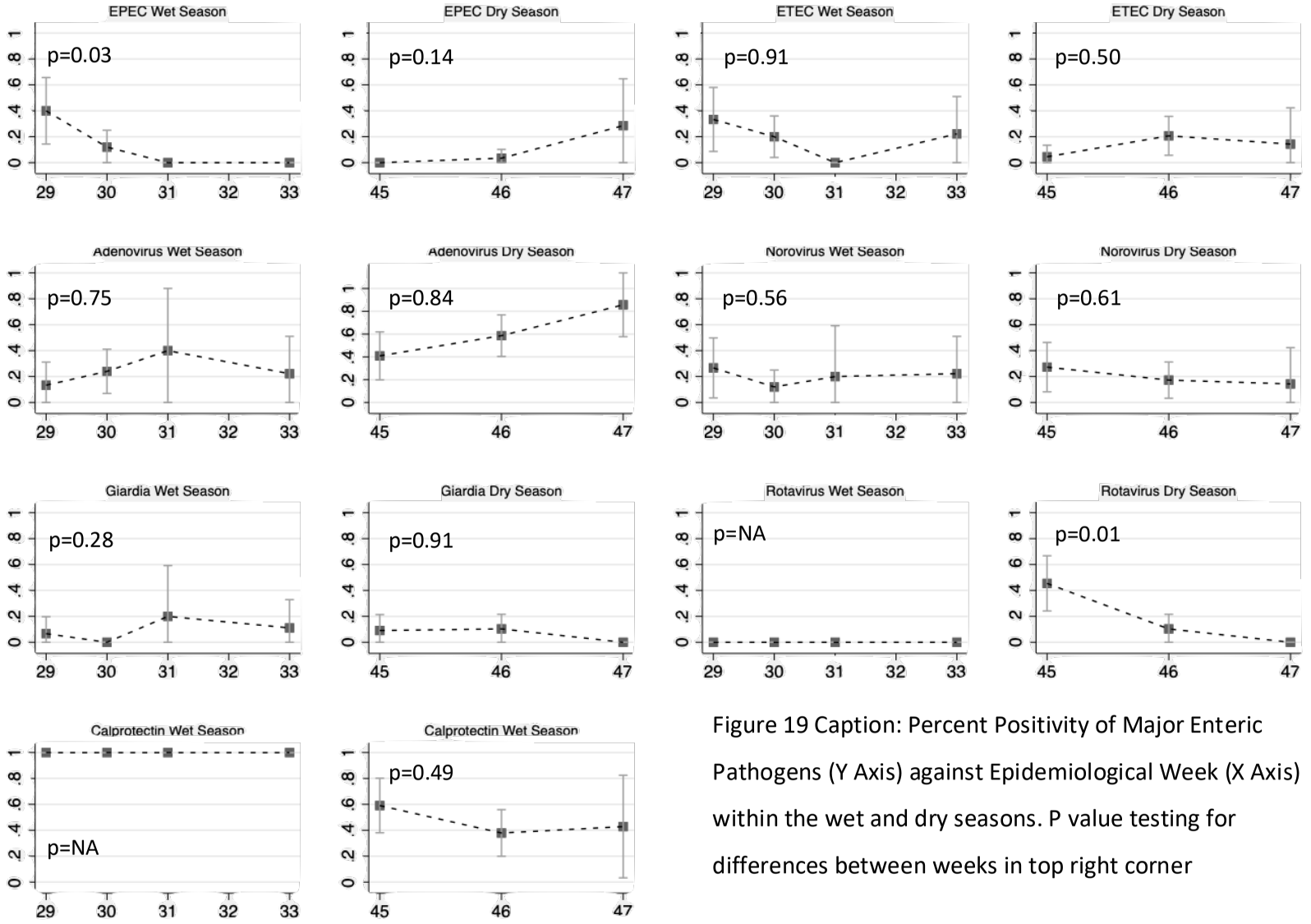


Figure 19 Caption: Percent Positivity of Major Enteric Pathogens (Y Axis) against Epidemiological Week (X Axis) within the wet and dry seasons. P value testing for differences between weeks in top right corner

Spatial Distribution

I found no evidence of spatial clustering for any pathogens, nor elevated proteins levels (Fig. 20). All values I obtained of Moran's I are close to 0, indicating random spatial distribution, albeit with generally a high degree of uncertainty (Table 19). There is additionally little evidence that distance between the home and water source or latrine is associated with increased risk of infection from EPEC, ETEC, Adenovirus, Norovirus, or Rotavirus, nor for elevated protein levels (Fig. 21 and 22). However, I do see evidence that the risk of Giardia is positively associated with increased distances between the home and both latrines and water sources, but with some uncertainty (Figures 21 and 22).

Temporal Aspects of Water and Sanitation Practices

My results on the spatio-temporal trends in enteric infection provide evidence that different enteric infections differ by season, but not spatially (at least in a small area such as the Leda Makeshift Refugee Camp), or temporally within seasons on the scale observed of 3-5 weeks (Table 17). While this seasonality is likely due in large part to changes in weather conditions between seasons, differences in WASH usage practices between seasons may also contribute to seasonality in enteric infection. When accessing water, respondents on average reported waiting substantially longer during the wet season (83 [71, 96] minutes) than the dry season (35 [27, 43] minutes). Further, fewer respondents reported doing something to make their water safe in the wet season (27% [17, 40]) than the dry season (65% [53, 76]). Similarly, fewer respondents reported washing their hands with water and soap in the wet season (84% [72,91]) than the dry season (97% [89,100]). The number that washed their hands at critical times (before preparing food, before eating, before feeding a child, after cleaning a child, and after defecating), however, was fairly similar between seasons. In

both seasons, a large number of respondents felt unsafe using the toilet both during the day and night.

Figure 20 Spatial Spread of Enteric Pathogens and Elevated Protein Levels Detected in Stool, by Week

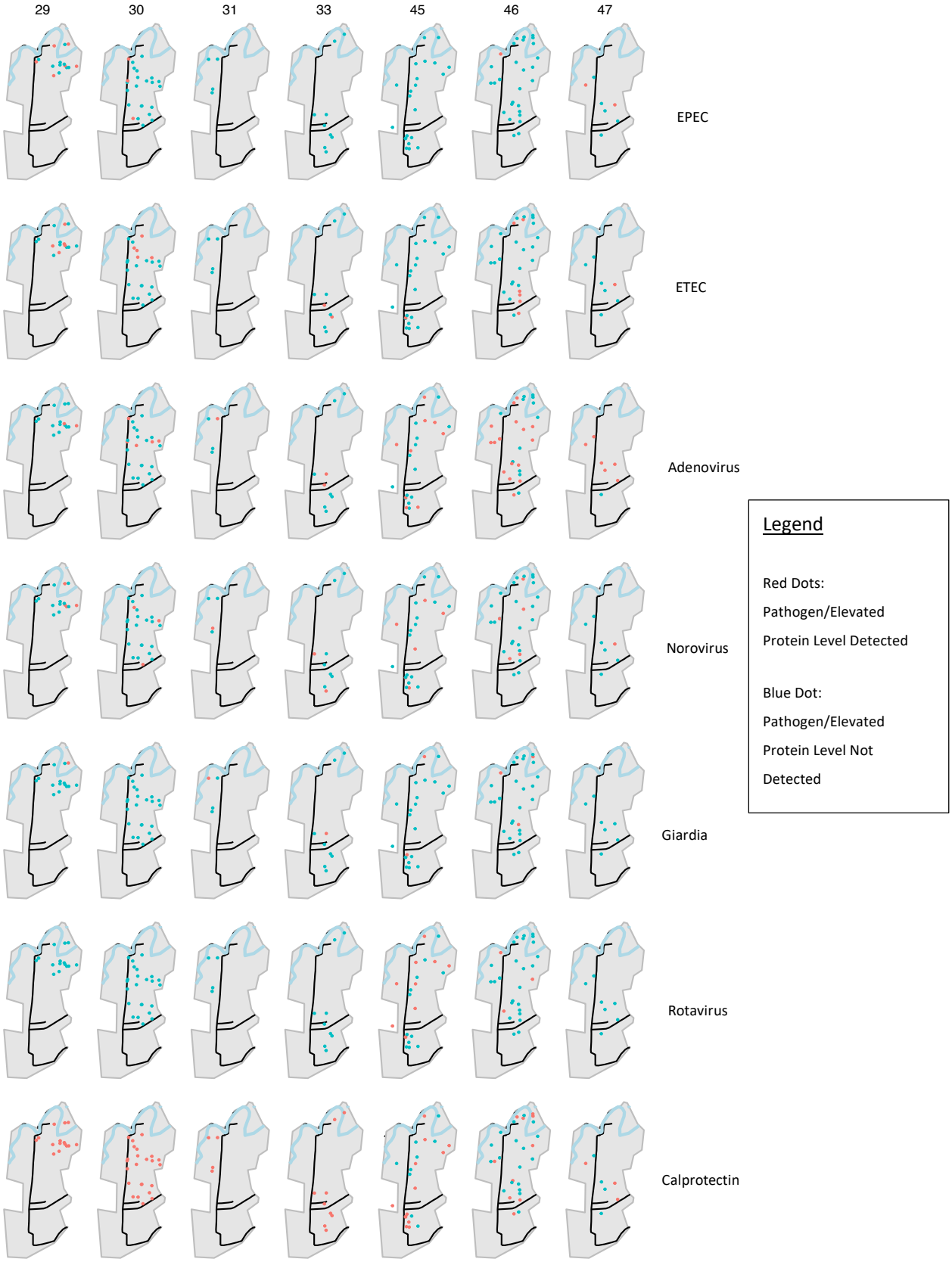


Table 19 Moran's I within Seasons

	Round 1 (Wet)	**Round 2 (Dry)**
EPEC	0.0 (p=0.72)	0.0 (p=0.74)
ETEC	0.0 (p=0.09)	0.0 (p=0.24)
Adenovirus	0.0 (p=0.34)	0.0 (p=0.15)
Norovirus	0.0 (p=0.68)	0.0 (p=0.61)
Giardia	0.0 (p=0.90)	0.0 (p=0.00)
Rotavirus	NA	0.0 (p=0.92)
Calprotectin	NA	0.0 (p=0.17)

Figure 21 Prevalence of Common Pathogens against Distance to Nearest Latrine

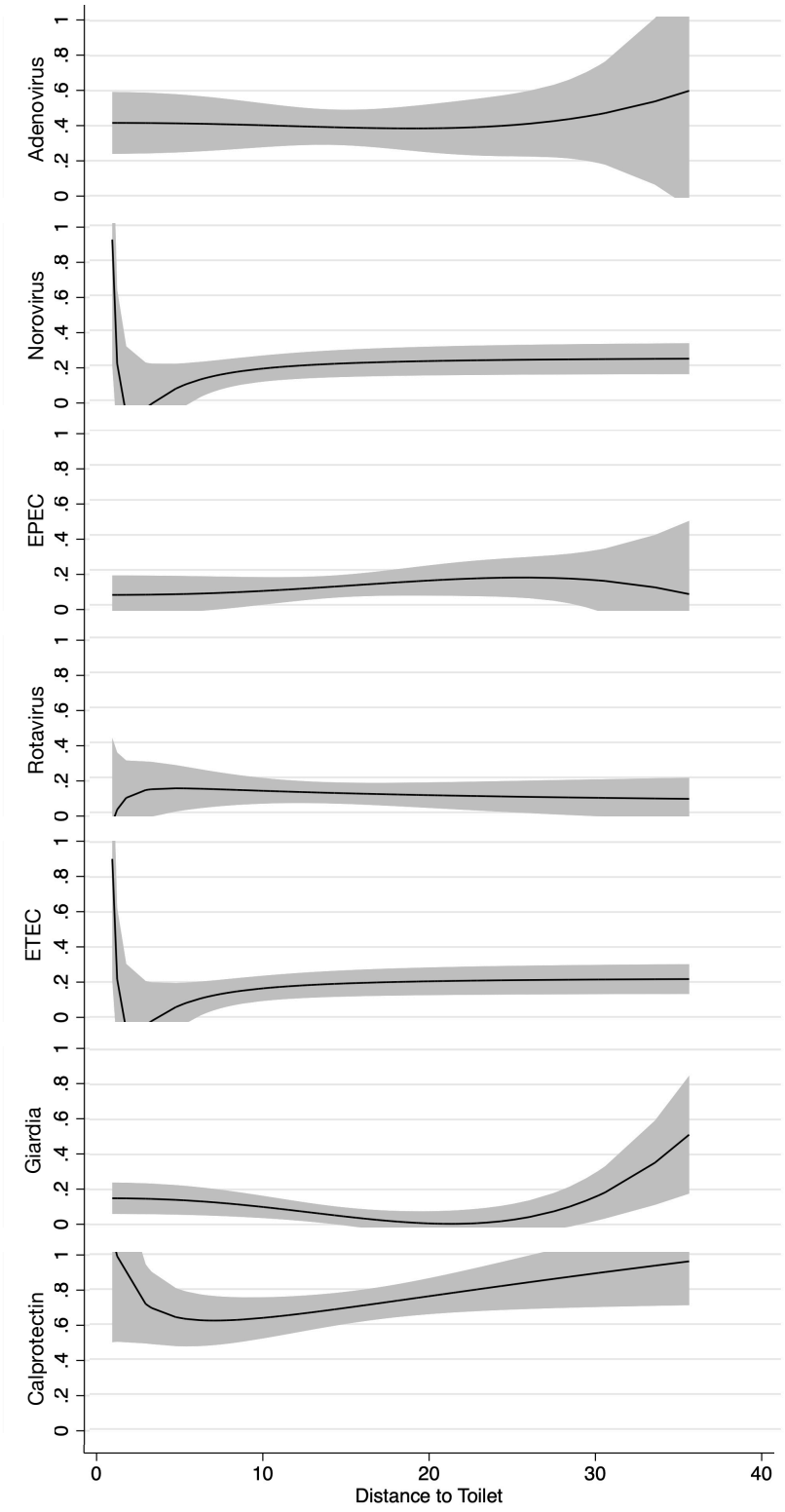
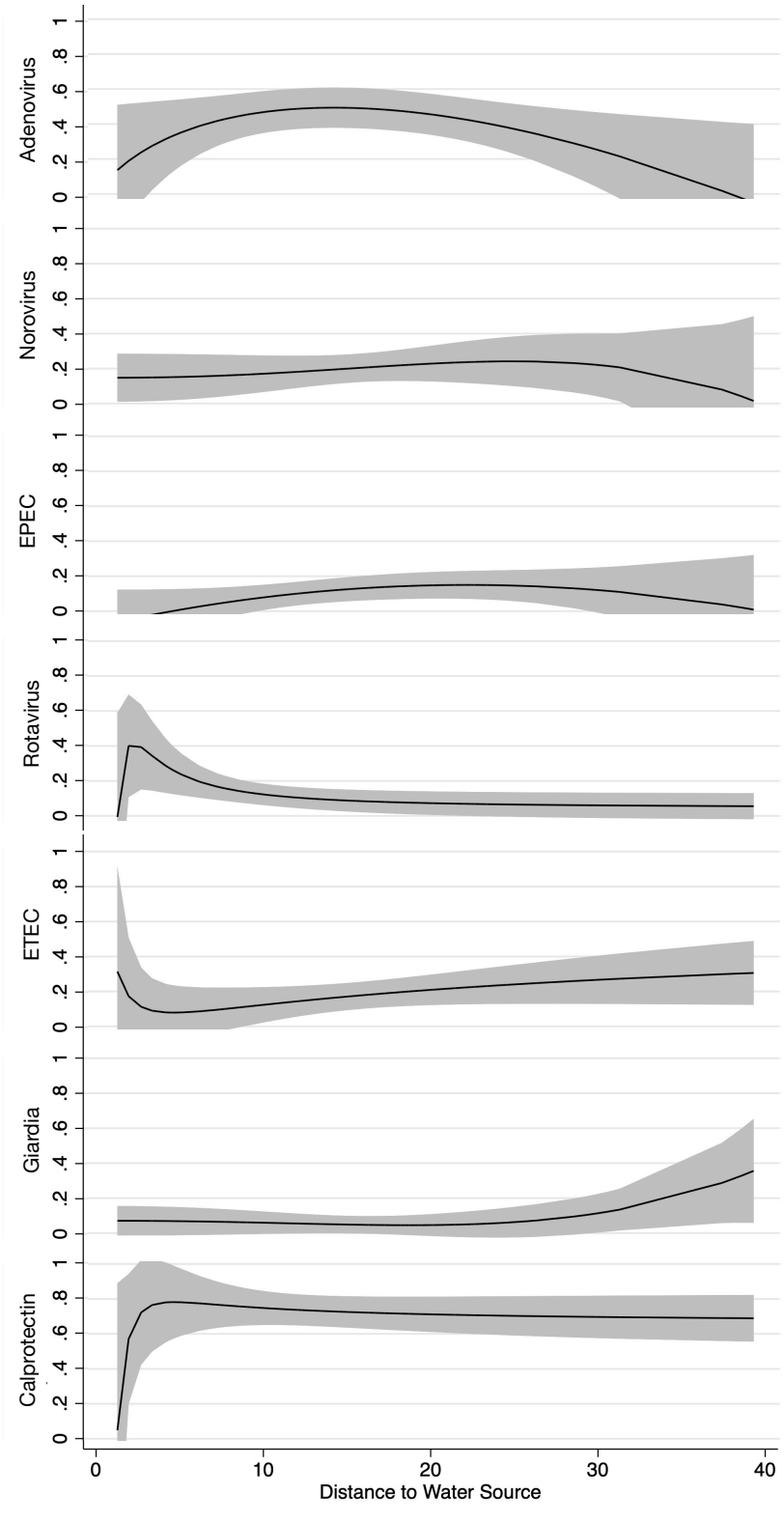


Figure 22 Prevalence of Common Pathogens against Distance to Nearest Water Source



5.5.Discussion

Summary of Findings

I found that the following enteric pathogens are present in the stools of under-fives living in the Leda Makeshift Camp area of the Cox's Bazar FDMN Camp: STEC, EPEC, ETEC, Salmonella, Campylobacter, Shigella, Cholera, Giardia, Rotavirus, Sapovirus, Adenovirus, and Norovirus. Past research has also found these pathogens to be present in the wider population of Bangladesh. For example, Qadri et al. (2007) reported that ETEC was found in 18% of stools in Dhaka after the 2004 Dhaka flooding, in line with my overall finding of 22.8% estimated prevalence in the rainy season[111]. Haque et al. (2003) reported a 3.3% estimated prevalence of EPEC in a Dhaka informal settlement, lower than my overall finding of 10.2% (averaged between seasons)[112]. Albert et al. (1995) however estimated an ETEC rate in Dhaka closer to ours, at 12.0%[113]. Dey et al. (2008) found evidence for Adenovirus in Dhaka, also at a lower rate than we did (2.8% compared to my 35%)[114]. This possibly indicates that there is a higher risk of Adenovirus in Cox's Bazar than Dhaka[114].

I additionally found evidence of seasonality in the prevalence of different types of enteric infection – viruses being more common in the dry season than the wet season, and the opposite for bacteria. My findings that elevated calprotectin levels are more common in the wet season than the dry season further support seasonality in bacterial infection as calprotectin levels are closely related to recent and present bacterial enteric infection[50].

My findings of seasonality in Cox's Bazar are consistent with previous research. Examining seasonality of viruses, Omore et al. (2016) estimated that Rotavirus is 10 percentage points more common in the dry season than the wet season in Western Kenya[109]. Berendes et al. (2017) estimated that there is a 40% lower risk of viral infection in the rainy season compared to the dry in Vellore, India[115]. Examining seasonality of bacteria, Feasey et al. (2015) found in Malawi that seasonal rainfall

was associated with an increase in Salmonella infection rates, although this may have been indirect as rainfall is associated with an increased risk of malnutrition[116]. Freeman et al. (2009) also report that gram-negative bacteria, such as EPEC and ETEC, are more common during wet seasons than dry seasons[117]. Similarly, Behiry et al. (2011) report that in Egypt, EPEC, as well as E. Coli in general, is more prevalent in the hot summer than cold winter[118].

We found no evidence that infections were clustered nor associated with household proximity to latrines or water taps. This suggests that the risk of infection is relatively uniform with respect to location in the camp. As such, management of these infections should be applied to the whole camp. However, the camp area which I examined was relatively small – the furthest distance between any two houses being 500 meters, which may partly explain this finding (along side some possible limitations, discussed below). In fact, past 'highly targeted' interventions against cholera outbreaks, which aim to work in small areas where risk is evenly distributed, have had intervention radii as large as 500 meters[119]. While I did not find many cases of cholera, it is likely that this is similar for outbreaks of other enteric pathogens. Spatial distribution would likely be apparent at larger scales, particularly if spaces are broken up by features that do not allow for population movement, such as large roads, rivers, or military blockades. I also for the most part did not find any association between infection risk and household distance from water taps and latrines. The small scale of the study (as well as the below discussed limitations) could again have contributed to this, with much larger distances from water taps and latrines needed to see a change in associated infection risk. Alternatively, it is possible that households and WASH facilities are not the primary source of infection for children, who spend most of their day elsewhere. Aside from evenings and mealtimes children play in communal "courtyards" near to the home, which are also often close to water and sanitation facilities. However, geospatial analyses in humanitarian settings are

scarce and further evidence is required to identify the primary sources of infection[120]. I later discuss possible future directions for this research.

Strengths and Limitations

As an analysis of secondary data, there are three substantial weaknesses in this study. Firstly, owing to budgetary constraints in the original study, there was a low sample size. While the sample size allowed for enquiry into the original research question in Chapter Four, it precluded some analyses into spatio-temporal risks, including adjusted multivariable analyses. Secondly, as the primary study was not designed to analyse the spatio-temporal risk of enteric infection, several important data points were not collected (such as where the child spends most of their day, or where the family obtains their food and water). Finally, due to logistical constraints, I was unable to visit homes throughout the entirety of the camp every week in round one. It is possible that homes in the areas I did not visit had a high level of infection, which would represent clustering which I had missed.

Alongside the limitations, there are substantial strengths in this work. We tested for a wide variety of pathogens, cutting down the chance that we omitted any important ones. Secondly, we also tested for calprotectin, a protein associated with recent infection (particularly bacterial infection), further reducing the chance that we missed infection by excluded pathogens[50]. We also tested our stool samples through PCR, the gold standard test for the presence of pathogens. Finally, our findings, particularly those on seasonality and the most common pathogens, are in agreement with current literature – including those discussed above in Bangladesh and other similar environments.

Future Research Directions

Further geospatial research in humanitarian settings is urgently needed. This research should build off our work by increasing sample size, looking at possible confounders or explanatory variables, and looking at a larger area of the camp of interest. Depending on the environment and epidemiology of enteric pathogens,

future research may choose to focus on fewer, more targeted pathogens than those we chose. Future research should also look at the spatial impact of WASH interventions, which would be of tremendous benefit to policy and program makers. Similarly, future research should look at the impact of topography and geography, as well as barriers to population movement (such as roads, rivers, and military blockades) on the spatial epidemiology of enteric pathogens.

Implications

My results, as well as those to follow from subsequent research, have implications for the surveillance and control of enteric pathogens in the Leda Makeshift FDMN Camp area in Cox's Bazar, and possibly in similar environments globally. As the profile of enteric infections present in the camp is similar to what has been found in the wider Bangladeshi population, I recommend that routine surveillance of enteric infections, and interventions to control enteric infections, do not discriminate between FDMNs and the local community, targeting the same pathogens in both. However, it is important to note the temporal differences that I found. There should be an emphasis on the surveillance, control, and treatment of bacterial disease during the wet season, and an emphasis on viral disease in the dry season, to best use the resources available. This can include surveillance (at a population level) for primarily viruses in the dry season and primarily bacteria in the wet season; clinicians testing patients who present to them with diarrhoea initially for the more common pathogens for the season; and timing mass vaccination programs by the season.

6. Conclusion

6.1.Summary of Key Findings

In this thesis, I aimed to examine the variability in estimated diarrhoea rates introduced by the methods used, and the extent of misclassification when using diarrhoea as a proxy marker of enteric infection; focusing on urban informal settlements and refugee camps. Chapters Two, Three, and Four demonstrate that the choice of method used to estimate diarrhoea rates can introduce variability into estimates, and that these different methods also do not agree with each other. Chapters Four and Five demonstrate the high extent of misclassification when using diarrhoea as a proxy marker of enteric infection, and the utility of measuring enteric infection directly.

Variability Introduced by Current Diarrhoea Rate Estimation Methods

In Chapter Two, through a systematic review and meta-analysis of the literature, I found that the two most common methods of diarrhoea rate estimation are passive surveillance (measurement of hospital cases for diarrhoea) and active surveillance (door to door surveying of diarrhoea). I found that passive surveillance is associated with lower estimated diarrhoea rates than active surveillance, possibly due to issues in accessing healthcare or choosing not to seek healthcare. These findings are in line with past studies which have found that: 1) people generally do not visit health facilities unless their child has severe diarrhoea, only 2% of cases; and 2) refugees and slum-dwellers face significant challenges in accessing healthcare[23, 27, 28 29-31].

In Chapters Two and Three, I also found evidence that factors specific to active surveillance impact estimated diarrhoea rates. Both chapters found evidence for respondent fatigue – getting tired of answering questions. In Chapter Two, this came in the form of frequent questioning resulting in lower estimated diarrhoea rates. In Chapter Three, a diarrhoea text messaging survey, this came in the form of asking more questions, particularly when asking more often, resulting in lower response rates. This is a new finding which has not been previously examined.

Both chapters additionally found evidence for recall bias, forgetting diarrhoea which happened in the past - finding that estimated diarrhoea rates are lower when you ask people to recall diarrhoea over prolonged periods, which has also been found in the past[33, 62].

In Chapter Four, through an observational comparison of four methods of diarrhoea measurement, I demonstrated that there is no agreement between the UNICEF/DHS method of active surveillance (asking a carer if their child had diarrhoea in the past 14 days), and two arguably more objective alternatives (asking carers to select pictures of the stool that their child had, and visually examining stool). No previous studies have examined the agreement between different measures of diarrhoea. However, this lack of agreement does support previous findings that carers are unable to reliably understand what is or is not diarrhoea, known as perception bias[37].

Together, my results, along with the existing evidence, indicate that current methods of diarrhoea rate estimation introduce bias and error. As such, these methods may not be fit for purpose. In epidemiological surveys when diarrhoea, not enteric infection, is of interest more objective methods should be used. Ideally, this would take the form of stool sampling for visual analysis as I examined in Chapter Four, and has also been examined in past research[37]. However, I do understand that this is demanding on the time of field enumerators, and as such not always possible. As an alternative, active surveillance with: 1) short recall periods (under 72 hours); 2) clear and simple questioning; and 3) questioning cross-sectionally (or, if longitudinally, using a long follow up period) should be used. If the context allows, I do recommend the use of text messaging for active surveillance due to increased privacy and possibly lower costs.

<u>Misclassification when using Diarrhoea as a Proxy Marker of Enteric Infection</u>

While it may be possible to reduce variation in estimated diarrhoea rates by using alternative methods, it is equally important to consider the extent of

misclassification when using diarrhoea as a proxy marker of enteric infection. Chapters Four examines this, revealing that in Cox's Bazar (and likely other similar environments as well) diarrhoea, through multiple measurement methods, offers no better indication of infection than a coin flip. The misclassification of enteric infection when using diarrhoea as a proxy marker is due to asymptomatic infection and non-pathogenic diarrhoea. While no studies have in the past examined diarrhoea as a proxy marker of enteric infection, past studies have found high rates of asymptomatic carriage in infants in LMICs – in line with my results[38].

In Chapter Five I demonstrate the utility of directly measuring enteric infection as an epidemiological tool. I show that different enteric infection causing pathogens have different risk factors and epidemiologies. As different enteric pathogens have different control measures, it is important that program makers understand which pathogens are endemic. This is not possible through surveillance for diarrhoea alone.

Together, Chapters Four and Five call for direct estimation of enteric infection, not diarrhoea, in epidemiological surveys, including surveillance activities and WASH trials. Direct estimation of infection allows for early detection of outbreaks, more accurate evaluations of WASH trials, and enhanced data for program makers. Considering my results from earlier chapters calling for less subjective methods in the measurement of gastrointestinal health, direct estimation of infection is by far the least subjective method possible and should be preferred.

6.2.External Validity

Confirming findings from my empirical studies, my systematic review revealed that the current methods used to estimate diarrhoea rates can introduce variation in estimates in every global region, and in both rural and urban geographies. However, it is important to acknowledge some limitations in this. Most of the studies included in my review were active surveillance studies. While active surveillance studies represent the majority of studies on population level

epidemiology and on the effectiveness of interventions designed to limit the spread of diarrhoea (including WASH interventions), it is likely that passive surveillance is more often used for routine population-level surveillance, which is generally unpublished. If I were able to include these, the results may be altered. Further, I was unable to assess the levels of reactivity in population-level diarrhoea measurement, which is likely to be a major contributing factor to bias in diarrhoea measurement. Finally, due to publication bias, the results may not adequately represent under-studied populations, such as internally displaced persons and those from certain nations.

To examine the extent of misclassification when using diarrhoea as a marker of enteric infection, I conducted an empirical study in the Cox's Bazar FDMN camp. It is possible that the results from this study suffer from external validity issues due to refugees acting differently to other groups; differences in endemic pathogen profiles around the world; and differences based on the environment, such as urban vs rural or climate. Further, it is possible that I did not include some disease causing pathogens in my analysis, either due to omission when determining which pathogens to test, or due to the inability to test for certain pathogens given COVID-19. However, the findings of this study are supported by other studies which found that enteric infection can have a high asymptomatic rate, and that diarrhoea can have non-pathogenic causes[38, 40, 100, 101].

6.3.Policy and Research Recommendations

My results have demonstrated that current methods of estimating diarrhoea rates can bias and introduce error to results; and that diarrhoea is not a suitable proxy of enteric infection in WASH trials and surveillance activities. Improvements to policy, as well as future research, is needed to combat these issues

<u>Improving the Measurement of Diarrhoea when only Disease is of Interest</u>

Sometimes the surveillance of diarrhoea without regard for infection status is well warranted. The DHS and MICS surveys are a prime example of this, as they do not

aim to discriminate by the cause of diarrhoea or establish levels of infection. In these situations, it is important to ensure that there is minimal bias and error in measurements of diarrhoea. As shown in chapters two and three, this would include having as short a recall period as possible (one or two days), questioning infrequently, and perhaps measuring prospectively rather than retrospectively. Future research should look into establishing the most cost effective methods: e.g. what is the maximum time period of recall without introducing recall bias; and how frequently can questioning be conducted without introducing respondent fatigue. Future research should also look into if parents can easily understand what is and is not diarrhoea, as we did in chapter four; and also examine the possible advantages of prospective questioning. If indeed parents are unable to understand what is diarrhoea, stool samples can be collected for visual analysis if resources allow. Further research should also be conducted into the use of technology such as SMS messaging and social media in diarrhoea surveillance (as well as the surveillance of other diseases).

Improving the Measurement of Enteric Infection in WASH Trials and Surveillance Activities

For many uses of diarrhoea measurement, such as WASH trials and outbreak diarrhoea is used only as a proxy for infection. My results have demonstrated that diarrhoea, measured through multiple methods, is neither a sensitive nor specific proxy marker of infection. As such, where infection is of interest, efforts should be taken to directly measure infection. The ideal method for this is the collection of stool samples for PCR-analysis of target enteric pathogens. However, I acknowledge that this may not be feasible in many settings, due to: 1) cost of PCR tests; 2) a lack of laboratory infrastructure; and 3) cost of sample collection. There are two possible solutions to these issues. Firstly, for low prevalence pathogens where a small increase in incidence is enough to trigger an outbreak (such as Cholera), pool testing may be used – where samples are combined and tested as one. This is not likely to work for WASH interventions however, as pathogens are likely to have a

high level of prevalence in areas where WASH interventions target. Similar to pooling of samples, faecal sludge can be tested (however more research is needed into testing of faecal sludge and sample pooling)[24]. However, a key limitation in this is the inability to discriminate infection by demographics, including importantly age. Secondly, for higher prevalence pathogens rapid diagnostic tests may be used. This would be particularly useful for evaluations of WASH interventions. It is even possible that participants would be able to test and report their own stool samples through a citizen science approach. However, the performance of rapid diagnostic tests for asymptomatic cases; the ability of beneficiaries to test their own stool; and lack of rapid diagnostic tests for many enteric pathogens remains an issue – calling for further research into this method.

7. Bibliography

1. (WHO) WHO. Diarrhoea Disease. World Health Organization. https://www.who.int/news-room/fact-sheets/detail/diarrhoeal-disease

2. Olack B, Feikin DR, Cosmas LO, et al. Mortality trends observed in population-based surveillance of an urban slum settlement, Kibera, Kenya, 2007–2010. *PloS one*. 2014;9(1):e85913.

3. Checkley W, Buckley G, Gilman RH, et al. Multi-country analysis of the effects of diarrhoea on childhood stunting. *International journal of epidemiology*. 2008;37(4):816-830.

4. Guerrant RL, Oria RB, Moore SR, Oria MO, Lima AA. Malnutrition as an enteric infectious disease with long-term effects on child development. *Nutr Rev*. Sep 2008;66(9):487-505. doi:10.1111/j.1753-4887.2008.00082.x

5. Scherzer AL, Chhagan M, Kauchali S, Susser E. Global perspective on early diagnosis and intervention for children with developmental delays and disabilities. *Developmental Medicine & Child Neurology*. 2012;54(12):1079-1084.

6. Ferdous F, Das SK, Ahmed S, et al. Diarrhoea in slum children: observation from a large diarrhoeal disease hospital in Dhaka, Bangladesh. *Tropical Medicine & International Health*. 2014;19(10):1170-1176. doi:https://doi.org/10.1111/tmi.12357

7. Peters DH, Garg A, Bloom G, Walker DG, Brieger WR, Hafizur Rahman M. Poverty and access to health care in developing countries. *Annals of the New York Academy of Sciences*. 2008;1136(1):161-171.

8. Chuah FLH, Tan ST, Yeo J, Legido-Quigley H. The health needs and access barriers among refugees and asylum-seekers in Malaysia: a qualitative study. *International journal for equity in health*. 2018;17(1):120-120. doi:10.1186/s12939-018-0833-x

9. Kyu HH, Shannon HS, Georgiades K, Boyle MH. Association of Urban Slum Residency with Infant Mortality and Child Stunting in Low and Middle Income Countries. *BioMed Research International*. 2013/09/17 2013;2013:604974. doi:10.1155/2013/604974

10. Garenne M. Urbanisation and child health in resource poor settings with special reference to under-five mortality in Africa. *Archives of Disease in Childhood*. 2010;95(6):464. doi:10.1136/adc.2009.172585

11. Friesen J, Taubenböck H, Wurm M, Pelz PF. Size distributions of slums across the globe using different data and classification methods. *European Journal of Remote Sensing*. 2019/08/09 2019;52(sup2):99-111. doi:10.1080/22797254.2019.1579617

12. UNHCR. Refugee Facts: Refugee Camps. UNHCR. 2020. https://www.unrefugees.org/refugee-facts/camps/

13. Friesen J, Rausch L, Pelz PF, Fürnkranz J. Determining factors for slum growth with predictive data mining methods. *Urban Science*. 2018;2(3):81.

14. Bonville C, Domachowske J. Cholera. In: Domachowske J, Suryadevara M, eds. *Vaccines: A Clinical Overview and Practical Guide*. Springer International Publishing; 2021:111-119.

15. WHO. The Use of Cholera Rapid Diagnostic Tests. 2021. https://www.who.int/cholera/task_force/Interim-guidance-cholera-RDT.pdf?ua=1

16. UNHCR. Emergency Handbook: Disease Surveillance Tresholds. 2.2 ed$_{2021}$.

17. Snow J. *On the mode of communication of cholera*. John Churchill; 1855.

18. Wolf J, Hunter PR, Freeman MC, et al. Impact of drinking water, sanitation and handwashing with soap on childhood diarrhoeal disease: updated meta-analysis and meta-regression. *Tropical medicine & international health*. 2018;23(5):508-525.

19. Pickering AJ, Null C, Winch PJ, et al. The WASH Benefits and SHINE trials: interpretation of WASH intervention effects on linear growth and diarrhoea. *Lancet Glob Health*. Aug 2019;7(8):e1139-e1146. doi:10.1016/S2214-109X(19)30268-2

20. Watson SI, Sartori J, Uthman O, Lilford RJ. Health effects of sanitation facilities: a Bayesian semiparametric analysis of compositional data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*. 2019;

21. Abolurin OO, Adegbola AJ, Oyelami OA, Adegoke SA, Bolaji OO. Vitamin A deficiency among under-five Nigerian children with diarrhoea. *African health sciences*. 2018;18(3):737-742.

22. Haug TT, Mykletun A, Dahl A. Are anxiety and depression related to gastrointestinal symptoms in the general population? *Scandinavian Journal of Gastroenterology*. 2002;37(3):294-298.

23. McNabb SJ, Conde JM, Ferland L, et al. *Transforming Public Health Surveillance-E-Book: Proactive Measures for Prevention, Detection, and Response*. Elsevier; 2016.

24. Capone D, Berendes D, Cumming O, et al. Analysis of Fecal Sludges Reveals Common Enteric Pathogens in Urban Maputo, Mozambique. *Environmental Science & Technology Letters*. 2020/10/13 2020;doi:10.1021/acs.estlett.0c00610

25. Babu A, Niehaus E. Multidisciplinary development of a proposed early warning and automated response system (EWARS) for epidemic prevention. Citeseer; 2010:

26. England PH. Notifiable diseases and causative organisms: how to report. London: HM Government; 2010.

27. Troeger C, Blacker BF, Khalil IA, et al. Estimates of the global, regional, and national morbidity, mortality, and aetiologies of diarrhoea in 195 countries: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet Infectious Diseases*. 2018;18(11):1211-1228.

28. Thiam S, Diene AN, Sy I, et al. Association between Childhood Diarrhoeal Incidence and Climatic Factors in Urban and Rural Settings in the Health District of Mbour, Senegal. *International journal of environmental research and public health*. 2017-1-1 2017;14(9)

29. Azage M, Kumie A, Worku A, C Bagtzoglou A, Anagnostou E. Effect of climatic variability on childhood diarrhea and its high risk periods in northwestern parts of Ethiopia. *PloS one*. 2017-1-1 2017;12(10):e0186933.

30. Norredam M. Migrants' access to healthcare. *Danish medical bulletin*. 2011;58(10):B4339-B4339.

31. Hjortsberg C. Why do the sick not utilise health care? The case of Zambia. *Health economics*. 2003;12(9):755-770.

32. Nelson AM, Perriëns JH, Kapita B, et al. A clinical and pathological comparison of the WHO and CDC case definitions for AIDS in Kinshasa, Zaïre: is passive surveillance valid? *AIDS*. 1993;7(9)

33. Feikin DR, Audi A, Olack B, et al. Evaluation of the optimal recall period for disease symptoms in home-based morbidity surveillance in rural and urban Kenya. *International journal of epidemiology*. 2010;39(2):450-458.

34. Zafar SN, Luby SP, Mendoza C. Recall errors in a weekly survey of diarrhoea in Guatemala: determining the optimal length of recall. 2010-1-1 %J Epidemiology and infection 2010;138(2):264-9.

35. Acharya D, Singh JK, Adhikari M, et al. Association of water handling and child feeding practice with childhood diarrhoea in rural community of Southern Nepal. *Journal of infection and public health*. 2018-1-1 2018;11(1):69-74.

36. Luby SP, Rahman M, Arnold BF, et al. Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Bangladesh: a cluster randomised controlled trial. 2018-1-1 %J The Lancet. Global health 2018;6(3):e302-e315.

37. Voskuijl W, Potani I, B, et al. Stool frequency recording in severe acute malnutrition ('StoolSAM'); an agreement study comparing maternal recall versus direct observation using diapers. 2017-1-1 %J BMC pediatrics 2017;17(1):140.

38. Okitsu S, Khamrin P, Takanashi S, et al. Molecular detection of enteric viruses in the stool samples of children without diarrhea in Bangladesh. *Infection, Genetics and Evolution*. 2020/01/01/ 2020;77:104055. doi:https://doi.org/10.1016/j.meegid.2019.104055

39. George CM, Burrowes V, Perin J, et al. Enteric infections in young children are associated with environmental enteropathy and impaired growth. *Tropical Medicine & International Health*. 2018;23(1):26-33.

40. Clay PG, Crutchley RD. Noninfectious Diarrhea in HIV Seropositive Individuals: a Review of Prevalence Rates, Etiology, and Management in the Era of Combination Antiretroviral Therapy. *Infectious Diseases and Therapy*. 2014/12/01 2014;3(2):103-122. doi:10.1007/s40121-014-0047-5

41. Sondaal SFV, Browne JL, Amoakoh-Coleman M, et al. Assessing the Effect of mHealth Interventions in Improving Maternal and Neonatal Care in Low- and Middle-Income Countries: A Systematic Review. *PLOS ONE*. 2016;11(5):e0154664. doi:10.1371/journal.pone.0154664

42. Silver L, Johnson C. Internet Connectivity Seen as Having Positive Impact on Life in Sub-Saharan Africa. Pew Research Center; 2018.

43. Pattnaik A, Mohan D, Chipokosa S, et al. Testing the validity and feasibility of using a mobile phone-based method to assess the strength of implementation of family planning programs in Malawi. *BMC health services research*. 2020;20(1):1-9.

44. Kuehne A, Lynch E, Marshall E, et al. Mortality, Morbidity and Health-Seeking Behaviour during the Ebola Epidemic 2014–2015 in Monrovia Results from a Mobile Phone Survey. *PLOS Neglected Tropical Diseases*. 2016;10(8):e0004899. doi:10.1371/journal.pntd.0004899

45. L'Engle K, Sefa E, Adimazoya EA, et al. Survey research with a random digit dial national mobile phone sample in Ghana: Methods and sample quality. *PloS one*. 2018;13(1):e0190902.

46. Vriesman MH, Velasco-Benitez CA, Ramirez CR, Benninga MA, Di Lorenzo C, Saps M. Assessing children's report of stool consistency: agreement between the pediatric Rome III questionnaire and the Bristol Stool Scale. *The Journal of Pediatrics*. 2017;190:69-73.

47. Sujatha B, Velayutham DR, Deivamani N, Bavanandam S. Normal bowel pattern in children and dietary and other precipitating factors in functional constipation. *Journal of clinical and diagnostic research: JCDR*. 2015;9(6):SC12.

48. Kotloff KL, Nataro JP, Blackwelder WC, et al. Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study. *Lancet*. Jul 20 2013;382(9888):209-22. doi:10.1016/s0140-6736(13)60844-2

49. Shastri YM, Bergis D, Povse N, et al. Prospective multicenter study evaluating fecal calprotectin in adult acute bacterial diarrhea. *The American journal of medicine*. 2008;121(12):1099-1106.

50. Chen C-C, Huang J-L, Chang C-J, Kong M-S. Fecal calprotectin as a correlative marker in clinical severity of infectious diarrhea and usefulness in evaluating bacterial or viral pathogens in children. *Journal of pediatric gastroenterology and nutrition*. 2012;55(5):541-547. doi:10.1097/MPG.0b013e318262a718

51. Chen C-C, Chang C-J, Lin T-Y, Lai M-W, Chao H-C, Kong M-S. Usefulness of fecal lactoferrin in predicting and monitoring the clinical severity of infectious diarrhea. *World journal of gastroenterology*. 2011;17(37):4218-4224. doi:10.3748/wjg.v17.i37.4218

52. UNOCHA. Rohingya Refugee Crisis. 2020.

53. Bhatia A, Mahmud A, Fuller A, et al. The Rohingya in Cox's Bazar: When the Stateless Seek Refuge. *Health and human rights*. 2018;20(2):105-122.

54. ODI. Case Study: Mwanza. London: Overseas Development Institute; 2016.

55. Group WB. World Bank Data. 2020.

56. Group WB. Tanzania's Path to Poverty Reduction and Pro-Poor Growth. 2019.

57. Canada IaRBo. Tanzania: Situation of Christians, including demographics, treatment and relations with other groups; state protection available in religious conflicts (2013-May 2016). 2016.

58. UNFPA. Tanzania: A Youthful and Rapidly Growing Nation.

59. Vasudevan L, Ostermann J, Moses SM, Ngadaya E, Mfinanga SG. Patterns of Mobile Phone Ownership and Use Among Pregnant Women in Southern Tanzania: Cross-Sectional Survey. *JMIR Mhealth Uhealth*. 2020/4/8 2020;8(4):e17122. doi:10.2196/17122

60. UN-Habitat. Youth-Led Mwanza City Informal Settlements Baseline Survey State of Living Conditions and Access to Urban Basic Services. 2018.

61. Tanzania Go. Mwanza Region Socio-Economic Profile. 1997.

62. Zafar SN, Luby SP, Mendoza C. Recall errors in a weekly survey of diarrhoea in Guatemala: determining the optimal length of recall. *Epidemiology and infection*. 2010-1-1 2010;138(2):264-9.

63. Larsen DA, Grisham T, Slawsky E, Narine L. An individual-level meta-analysis assessing the impact of community-level sanitation access on child stunting, anemia, and diarrhea: Evidence from DHS and MICS surveys. *PLOS Neglected Tropical Diseases*. 2017;11(6):e0005591. doi:10.1371/journal.pntd.0005591

64. Gizaw Z, Woldu W, Bitew BD. Child feeding practices and diarrheal disease among children less than two years of age of the nomadic people in Hadaleala District, Afar Region, Northeast Ethiopia. *International breastfeeding journal*. 2017;12(1):24.

65. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ*. 2009;339:b2535. doi:10.1136/bmj.b2535

66. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of Observational Studies in EpidemiologyA Proposal for Reporting. *JAMA*. 2000;283(15):2008-2012. doi:10.1001/jama.283.15.2008

67. Ryan R, Samuel W, Richard L. **Diarrhoea prevalence measurement methods in LMICs: a systematic review (Protocol)**. PROSPEREO2018.

68. Development OfEC-oa. DAC List of ODA Recipients. Organisation for Economic Co-operation and Development. 2019. https://www.oecd.org/dac/financing-sustainable-development/development-finance-standards/DAC_List_ODA_Recipients2018to2020_flows_En.pdf

69. Division UNS. SDG Indicators Regional groupings used in 2017 Report and Statistical Annex. United Nations. 2020. https://unstats.un.org/sdgs/indicators/regional-groups/

70. *StataSE Version 15*. StataCorp; 2017.

71. Zwane AP, Zinman J, Van Dusen E, et al. Being surveyed can change later behavior and related parameter estimates. *Proceedings of the National Academy of Sciences*. 2011;108(5):1821-1826.

72. Melo MCNd, Taddei JAdAC, Diniz-Santos DR, May DS, Carneiro NB, Silva LR. Incidence of diarrhea: poor parental recall ability. *The Brazilian journal of infectious diseases : an official publication of the Brazilian Society of Infectious Diseases*. 2007-1-1 2007;11(6):571-9.

73. Lee G, Cama V, Gilman RH, et al. Comparison of two types of epidemiological surveys aimed at collecting daily clinical symptoms in community-based longitudinal studies. 2010-1-1 %J Annals of epidemiology 2010;20(2):151-8.

74. Melo MCNd, Taddei JAdAC, Diniz-Santos DR, May DS, Carneiro NB, Silva LR. Incidence of diarrhea: poor parental recall ability. 2007-1-1 %J The Brazilian journal of infectious diseases : an official publication of the Brazilian Society of Infectious Diseases 2007;11(6):571-9.

75. Rego R, Watson S, Ishengoma P, Langat P, Otieno HP, Lilford R. Effectiveness of SMS messaging for diarrhoea measurement: a factorial cross-over randomised controlled trial. *BMC Medical Research Methodology*. 2020/06/30 2020;20(1):174. doi:10.1186/s12874-020-01062-3

76. Wood L, Egger M, Gluud LL, et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *BMJ*. 2008;336(7644):601. doi:10.1136/bmj.39465.451748.AD

77. Cumming O, Curtis V. Implications of WASH Benefits trials for water and sanitation. *Lancet Glob Health*. Jun 2018;6(6):e613-e614. doi:10.1016/s2214-109x(18)30192-x

78. Humphrey JH, Mbuya MNN, Ntozini R, et al. Independent and combined effects of improved water, sanitation, and hygiene, and improved complementary feeding, on child stunting and anaemia in rural Zimbabwe: a cluster-randomised trial. *Lancet Glob Health*. Jan 2019;7(1):e132-e147. doi:10.1016/S2214-109X(18)30374-7

79. Luby SP, Rahman M, Arnold BF, et al. Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Bangladesh: a cluster randomised controlled trial. *The Lancet Global Health*. 2018;6(3):e302-e315. doi:10.1016/S2214-109X(17)30490-4

80. Null C, Stewart CP, Pickering AJ, et al. Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Kenya: a cluster-randomised controlled trial. *The Lancet Global Health*. 2018;6(3):e316-e329. doi:10.1016/S2214-109X(18)30005-6

81. GREENLAND S. Basic Methods for Sensitivity Analysis of Biases. *International Journal of Epidemiology*. 1996;25(6):1107-1116. doi:10.1093/ije/25.6.1107-a

82. Mukund KB, Murray PJ. Cell phone short messaging service (SMS) for HIV/AIDS in South Africa: a literature review. *Studies in health technology and informatics*. 2010;160(Pt 1):530-534.

83. Schulz KF, Altman DG, Moher D, the CG. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *BMC Medicine*. 2010/03/24 2010;8(1):18. doi:10.1186/1741-7015-8-18

84. Montgomery AA, Peters TJ, Little P. Design, analysis and presentation of factorial randomised controlled trials. *BMC medical research methodology*. 2003;3(1):26.

85. Bhavnani D, Goldstick JE, Cevallos W, Trueba G, Eisenberg JNS. Impact of rainfall on diarrheal disease risk associated with unimproved water and sanitation. *The American journal of tropical medicine and hygiene*. 2014-1-1 2014;90(4):705-11.

86. Hopkins KD, Gullickson AR. Response Rates in Survey Research: A Meta-Analysis of the Effects of Monetary Gratuities. *The Journal of Experimental Education*. 1992/07/01 1992;61(1):52-62. doi:10.1080/00220973.1992.9943849

87. Fuller JA, Clasen T, Heijnen M, Eisenberg JNS. Shared sanitation and the prevalence of diarrhea in young children: evidence from 51 countries, 2001-2011. *The American journal of tropical medicine and hygiene*. 2014-1-1 2014;91(1):173-80.

88. Blum D, Feachem RG. Measuring the impact of water supply and sanitation investments on diarrhoeal diseases: problems of methodology. *International journal of epidemiology*. 1983;12(3):357-365.

89. Voskuijl W, Potani I, B, et al. Stool frequency recording in severe acute malnutrition ('StoolSAM'); an agreement study comparing maternal recall versus direct observation using diapers. *BMC pediatrics*. 2017-1-1 2017;17(1):140.

90. Thiam S, Diene AN, Sy I, et al. Association between Childhood Diarrhoeal Incidence and Climatic Factors in Urban and Rural Settings in the Health District of Mbour, Senegal. 2017-1-1 %J International journal of environmental research and public health 2017;14(9)

91. Bekkali N, Hamers SL, Reitsma JB, Van Toledo L, Benninga MA. Infant stool form scale: development and results. *The Journal of pediatrics*. 2009;154(4):521-526. e1. doi:10.1016/j.jpeds.2008.10.010

92. Force BA. CLIMATE OF COX'S BAZAR. 2018.

93. Chao L-W, Szrek H, Peltzer K, et al. A Comparison of EPI Sampling, Probability Sampling, and Compact Segment Sampling Methods for Micro and Small Enterprises. *Journal of development economics*. 2012;98(1):94-107. doi:10.1016/j.jdeveco.2011.08.007

94. Khan MN, Islam MM. Effect of exclusive breastfeeding on selected adverse health and nutritional outcomes: a nationally representative study. *BMC public health*. 2017-1-1 2017;17(1):889.

95. Bekkali N, Hamers SL, Reitsma JB, Van Toledo L, Benninga MA. Infant stool form scale: development and results. *J Pediatr*. Apr 2009;154(4):521-526.e1. doi:10.1016/j.jpeds.2008.10.010

96. Contol CfD. Stool Specimens - Molecular Diagnosis. https://www.cdc.gov/dpdx/diagnosticprocedures/stool/moleculardx.html

97. Seed P. DIAGT: Stata module to report summary statistics for diagnostic tests compared to true disease status. 2010;

98. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam med*. 2005;37(5):360-363.

99. Nyeko R, Kalyesubula I, Mworozi E, Bachou H. Lactose intolerance among severely malnourished children with diarrhoea admitted to the nutrition unit, Mulago hospital, Uganda. *BMC pediatrics*. 2010;10(1):31.

100. Bergogne-Bérézin E. Treatment and prevention of antibiotic associated diarrhea. *International Journal of Antimicrobial Agents*. 2000/12/01/ 2000;16(4):521-526. doi:https://doi.org/10.1016/S0924-8579(00)00293-4

101. Korpe PS, Petri WA, Jr. Environmental enteropathy: critical implications of a poorly understood condition. *Trends Mol Med*. Jun 2012;18(6):328-36. doi:10.1016/j.molmed.2012.04.007

102. Diggle PJ. Estimating Prevalence Using an Imperfect Test. *Epidemiology Research International*. 2011/10/23 2011;2011:608719. doi:10.1155/2011/608719

103. Oberhelman RA, Flores-Abuxapqui J, Suarez-Hoil G, et al. Asymptomatic salmonellosis among children in day-care centers in Merida, Yucatan, Mexico. 2001-1-1 %J The Pediatric infectious disease journal 2001;20(8):792-7.

104. Prado MS, Cairncross S, Strina A, Barreto ML, Oliveira-Assis AM, Rego S. Asymptomatic giardiasis and growth in young children; a longitudinal study in Salvador, Brazil. *Parasitology*. 2005-1-1 2005;131:51-6.

105. Hershey CL, Doocy S, Anderson J, Haskew C, Spiegel P, Moss WJ. Incidence and risk factors for Malaria, pneumonia and diarrhea in children under 5 in UNHCR refugee camps: A retrospective study. *Conflict and health*. 2011;5(1):24-24. doi:10.1186/1752-1505-5-24

106. Burnet E, Rudge JW. Diarrhoeal disease outbreaks associated with sanitation provision failures in refugee camps worldwide: a literature review. *Waterlines*. 2019;38(1):36-63.

107. Okoyo C, Nikolay B, Kihara J, et al. Monitoring the impact of a national school based deworming programme on soil-transmitted helminths in Kenya: the first three years, 2012–2014. *Parasites & vectors*. 2016;9(1):408.

108. Neilan RLM, Schaefer E, Gaff H, Fister KR, Lenhart S. Modeling optimal intervention strategies for cholera. *Bulletin of mathematical biology*. 2010;72(8):2004-2018.

109. Omore R, Tate JE, O'Reilly CE, et al. Epidemiology, seasonality and factors associated with rotavirus infection among children with moderate-to-severe diarrhea in rural western Kenya, 2008–2012: the Global Enteric Multicenter Study (GEMS). *PloS one*. 2016;11(8):e0160060.

110. Matthews JL, Diawara N, Waller LA. Quantifying Spatio-Temporal Characteristics via Moran's Statistics. In: Diawara N, ed. *Modern Statistical Methods for Spatial and Multivariate Data*. Springer International Publishing; 2019:163-177.

111. Qadri F, Saha A, Ahmed T, Al Tarique A, Begum YA, Svennerholm A-M. Disease burden due to enterotoxigenic Escherichia coli in the first 2 years of life in an urban community in Bangladesh. *Infection and immunity*. 2007-1-1 2007;75(8):3961-8.

112. Haque R, Mondal D, Kirkpatrick BD, et al. Epidemiologic and clinical characteristics of acute diarrhea with emphasis on Entamoeba histolytica infections

in preschool children in an urban slum of Dhaka, Bangladesh. 2003-1-1 %J The American journal of tropical medicine and hygiene 2003;69(4):398-405.

113. Albert MJ, Faruque S, Faruque A, et al. Controlled study of Escherichia coli diarrheal infections in Bangladeshi children. *Journal of Clinical Microbiology*. 1995;33(4):973-977.

114. Dey SK, Shimizu H, Phan TG, et al. Molecular epidemiology of adenovirus infection among infants and children with acute gastroenteritis in Dhaka City, Bangladesh. *Infection, Genetics and Evolution*. 2009;9(4):518-522.

115. Berendes D, Leon J, Kirby A, et al. Household sanitation is associated with lower risk of bacterial and protozoal enteric infections, but not viral infections and diarrhoea, in a cohort study in a low-income urban neighbourhood in Vellore, India. 2017-1-1 %J Tropical medicine & international health : TM & IH 2017;22(9):1119-1129.

116. Feasey NA, Everett D, Faragher EB, et al. Modelling the contributions of malaria, HIV, malnutrition and rainfall to the decline in paediatric invasive non-typhoidal Salmonella disease in Malawi. *PLoS Negl Trop Dis*. 2015;9(7):e0003979.

117. Freeman J, Anderson D, Sexton D. Seasonal peaks in Escherichia coli infections: possible explanations and implications. *Clinical Microbiology and Infection*. 2009;15(10):951-953.

118. Behiry IK, Abada EA, Ahmed EA, Labeeb RS. Enteropathogenic Escherichia coli associated with diarrhea in children in Cairo, Egypt. *TheScientificWorldJOURNAL*. 2011;11

119. Ratnayake R, Finger F, Azman AS, et al. Highly targeted spatiotemporal interventions against cholera epidemics, 2000–19: a scoping review. *The Lancet Infectious Diseases*. 2020;

120. Jabbour S, Attal B. Geospatial analysis: a new frontier in humanitarian health research? *The Lancet Global Health*. 2020;8(11):e1353-e1354.

8.

Appendix I: Search String for the Systematic Review

1. diarrh?ea? or AWD or loose stool or watery stool or dysentery or ABD or fe?cal
2. Developing Country**.sh.**
3. (Africa or Asia or Caribbean or West Indies or South America or Latin America or Central America)**.hw,ti,ab,cp.**
4. (Afghanistan or Albania or Algeria or Angola or Antigua or Barbuda or Argentina or Armenia or Armenian or Aruba or Azerbaijan or Bahrain or Bangladesh or Barbados or Benin or Byelarus or Byelorussian or Belarus or Belorussian or Belorussia or Belize or Bhutan or Bolivia or Bosnia or Herzegovina or Hercegovina or Botswana or Brasil or Brazil or Bulgaria or Burkina Faso or Burkina Fasso or Upper Volta or Burundi or Urundi or Cambodia or Khmer Republic or Kampuchea or Cameroon or Cameroons or Cameron or Camerons or Cape Verde or Central African Republic or Chad or Chile or China or Colombia or Comoros or Comoro Islands or Comores or Mayotte or Congo or Zaire or Costa Rica or Cote d'Ivoire or Ivory Coast or Croatia or Cuba or Cyprus or Czechoslovakia or Czech Republic or Slovakia or Slovak Republic or Djibouti or French Somaliland or Dominica or Dominican Republic or East Timor or East Timur or Timor Leste or Ecuador or Egypt or United Arab Republic or El Salvador or Eritrea or Estonia or Ethiopia or Fiji or Gabon or Gabonese Republic or Gambia or Gaza or Georgia Republic or Georgian Republic or Ghana or Gold Coast or Greece or Grenada or Guatemala or Guinea or Guam or Guiana or Guyana or Haiti or Honduras or Hungary or India or Maldives or Indonesia or Iran or Iraq or Isle of Man or Jamaica or Jordan or Kazakhstan or Kazakh or Kenya or Kiribati or Korea or Kosovo or Kyrgyzstan or Kirghizia or Kyrgyz Republic or Kirghiz or Kirgizstan or Lao PDR or Laos or Latvia or Lebanon or Lesotho or Basutoland or Liberia or Libya or Lithuania or Macedonia or Madagascar or Malagasy Republic or Malaysia or Malaya or Malay or Sabah or Sarawak or Malawi or Nyasaland or Mali or Malta or Marshall Islands or Mauritania or Mauritius or Agalega Islands or Mexico or Micronesia or Middle East or Moldova or Moldovia or Moldovian or Mongolia or Montenegro or Morocco or Ifni or Mozambique or Myanmar or Myanma or Burma or Namibia or Nepal or Netherlands Antilles or New Caledonia or Nicaragua or Niger or Nigeria or Northern Mariana Islands or Oman or Muscat or Pakistan or Palau or Palestine or Panama or Paraguay or Peru or Philippines or Philipines or Phillipines or Phillippines or Poland or Portugal or Puerto Rico or Romania or Rumania or Roumania or Russia or Russian or Rwanda or Ruanda or Saint Kitts or St Kitts or Nevis or Saint Lucia or St Lucia or Saint Vincent or St Vincent or Grenadines or Samoa or Samoan Islands or Navigator Island or Navigator Islands or Sao Tome or Saudi Arabia or Senegal or Serbia or Montenegro or Seychelles or Sierra Leone or Slovenia or Sri Lanka or Ceylon or Solomon Islands or Somalia or South Africa or Sudan or Suriname or Surinam or

Swaziland or Syria or Tajikistan or Tadzhikistan or Tadjikistan or Tadzhik or Tanzania or Thailand or Togo or Togolese Republic or Tonga or Trinidad or Tobago or Tunisia or Turkey or Turkmenistan or Turkmen or Uganda or Ukraine or Uruguay or USSR or Soviet Union or Union of Soviet Socialist Republics or Uzbekistan or Uzbek or Vanuatu or New Hebrides or Venezuela or Vietnam or Viet Nam or West Bank or Yemen or Yugoslavia or Zambia or Zimbabwe or Rhodesia).**hw,ti,ab,cp.**

5. ((developing or less* developed or under developed or underdeveloped or middle income or low* income or underserved or under served or deprived or poor*) adj (countr* or nation? or population? or world))**.ti,ab.**
6. ((developing or less* developed or under developed or underdeveloped or middle income or low* income) adj (economy or economies))**.ti,ab.**
7. (low* adj (gdp or gnp or gross domestic or gross national))**.ti,ab.**
8. (low adj3 middle adj3 countr*)**.ti,ab.**
9. (lmic or lmics or third world or lami countr*)**.ti,ab.**
10. transitional countr***.ti,ab.**
11. or/1-9
12. measurement or burden or survey or questionnaire or stool sample or collection or (Bristol adj2 scale) or (Amsterdam adj2 scale) or scale or pathogen or microbiological or biological or protein or olfaction or prevalence or incidence or odds or risk
13. exp animals/ not humans.sh.
14. 1 and 11 and 12
15. 14 not 13
16. Limit to ((english or french) and yr="1993 -Current")

9. Appendix II: Studies Included in the Systematic Review

1 Miranda JJ, Davies AR, Smith GD, et al. Frequency of diarrhoea as a predictor of elevated blood pressure in children; 2009.

2 Panda S, Deb AK, Chawla-Sarkar M, et al. Factors associated with diarrhoea in young children and incidence of symptomatic rotavirus infection in rural West Bengal, India; 2014.

3 Abu Mourad TA. Palestinian refugee conditions associated with intestinal parasites and diarrhoea: Nuseirat refugee camp as a case study. *Public health* 2004; **118**(2): 131-42.

4 Abu-Elyazeed RR, Wierzba TF, Frenck RW, et al. Epidemiology of Shigella-associated diarrhea in rural Egyptian children. *The American journal of tropical medicine and hygiene* 2004; **71**(3): 367-72.

5 Acharya D, Singh JK, Adhikari M, et al. Association of water handling and child feeding practice with childhood diarrhoea in rural community of Southern Nepal. *Journal of infection and public health* 2018; **11**(1): 69-74.

6 Adane M, Mengistie B, Kloos H, Medhin G, Mulat W. Sanitation facilities, hygienic conditions, and prevalence of acute diarrhea among under-five children in slums of Addis Ababa, Ethiopia: Baseline survey of a longitudinal study. *PloS one* 2017; **12**(8): e0182783.

7 Agtini MD, Soeharno R, Lesmana M, et al. The burden of diarrhoea, shigellosis, and cholera in North Jakarta, Indonesia: findings from 24 months surveillance. *BMC infectious diseases* 2005; **5**: 89.

8 Agustina R, Sari TP, Satroamidjojo S, Bovee-Oudenhoven IMJ, Feskens EJM, Kok FJ. Association of food-hygiene practices and diarrhea prevalence among Indonesian

young children from low socioeconomic urban areas. *BMC public health* 2013; **13**: 977.

9 Agustina R, Shankar AV, Ayuningtyas A, Achadi EL, Shankar AH. Maternal agency influences the prevalence of diarrhea and acute respiratory tract infections among young Indonesian children. *Maternal and child health journal* 2015; **19**(5): 1033-46.

10 Ahiadeke C. Breast-feeding, diarrhoea and sanitation as components of infant and child health: a study of large scale survey data from Ghana and Nigeria. *Journal of biosocial science* 2000; **32**(1): 47-61.

11 Kandala N-B, Emina JB, Nzita PDK, Cappuccio FP. Diarrhoea, acute respiratory infection, and fever among children in the Democratic Republic of Congo; 2009.

12 Kandala N-B, Ji C, Stallard N, Stranges S, Cappuccio FP. Spatial analysis of risk factors for childhood morbidity in Nigeria; 2007.

13 Kandala N-B, Magadi MA, Madise NJ. An investigation of district spatial variations of childhood diarrhoea and fever morbidity in Malawi; 2006.

14 Alam DS, Marks GC, Baqui AH, Yunus M, Fuchs GJ. Association between clinical type of diarrhoea and growth of children under 5 years in rural Bangladesh. *International journal of epidemiology* 2000; **29**(5): 916-21.

15 Ali M, Atkinson D, Underwood P. Determinants of use rate of oral rehydration therapy for management of childhood diarrhoea in rural Bangladesh. *Journal of health, population, and nutrition* 2000; **18**(2): 103-8.

16 Aluisio AR, Maroof Z, C, et al. Risk factors associated with recurrent diarrheal illnesses among children in Kabul, Afghanistan: a prospective cohort study. *PloS one* 2015; **10**(2): e0116342.

17 Aluisio AR, Maroof Z, C, et al. Vitamin D3supplementation and childhood diarrhea: a randomized controlled trial. *Pediatrics* 2013; **132**(4): e832-40.

18 Aluko OO, Afolabi OT, Olaoye EA, Adebayo AD, Oyetola SO, Abegunde OO. The management of the faeces passed by under five children: an exploratory, cross-sectional research in an urban community in Southwest Nigeria. *BMC public health* 2017; **17**(1): 178.

19 Amugsi DA, Aborigo RA, Oduro AR, Asoala V, Awine T, Amenga-Etego L. Socio-demographic and environmental determinants of infectious disease morbidity in children under 5 years in Ghana. *Global health action* 2015; **8**: 29349.

20 Anders KL, Thompson CN, Thuy NTV, et al. The epidemiology and aetiology of diarrhoeal disease in infancy in southern Vietnam: a birth cohort study. *International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases* 2015; **35**: 3-10.

21 Anderson LC, Tegegn A, Tessema F, Galea S, r, Hadley C. Food insecurity, childhood illness and maternal emotional distress in Ethiopia. *Public health nutrition* 2012; **15**(4): 648-55.

22 Anteneh ZA, Andargie K, Tarekegn M. Prevalence and determinants of acute diarrhea among children younger than five years old in Jabithennan District, Northwest Ethiopia, 2014. *BMC public health* 2017; **17**(1): 99.

23 Bhandari N, Bahl R, Mazumdar S, et al. Effect of community-based promotion of exclusive breastfeeding on diarrhoeal illness and growth: a cluster randomised controlled trial; 2003.

24 Bhandari N, Bahl R, Taneja S, et al. Substantial reduction in severe diarrheal morbidity by daily zinc supplementation in young north Indian children; 2002.

25 Arnold BF, Khush RS, Ramaswamy P, et al. A causal framework for evaluating pre-existing interventions: An example motivated by efforts in the water, sanitation and hygiene sector in rural India. *American Journal of Tropical Medicine and Hygiene* 2010; **83**(5): 7.

26 Arnold B, Arana B, Mausezahl D, Hubbard A, Colford JM, Jr. Evaluation of a pre-existing, 3-year household water treatment and handwashing intervention in rural Guatemala. *International journal of epidemiology* 2009; **38**(6): 1651-61.

27 Arrizon AV, Andersson N, Ledogar RJ. Micro-regional planning: evidence-based community buy-in for health development in five of Mexico's poorest rural districts. *BMC health services research* 2011; **11**: S2.

28 Asling-Monemi K, Naved RT, Persson LA. Violence against women and increases in the risk of diarrheal disease and respiratory tract infections in infancy: a prospective cohort study in Bangladesh. *Archives of pediatrics & adolescent medicine* 2009; **163**(10): 931-6.

29 Assis AMO, Barreto ML, Santos LMP, Fiaccone R, da Silva Gomes GS. Growth faltering in childhood related to diarrhea: a longitudinal community based study. *European journal of clinical nutrition* 2005; **59**(11): 1317-23.

30 Aung T, M, W, Khin HSS, Montagu D. Incidence of pediatric diarrhea and public-private preferences for treatment in rural Myanmar: a randomized cluster survey. *Journal of tropical pediatrics* 2013; **59**(1): 10-6.

31 Aung T, Montagu D, Su Su Khin H, et al. Impact of a social franchising program on uptake of oral rehydration solution plus zinc for childhood diarrhea in myanmar: a community-level randomized controlled trial. *Journal of tropical pediatrics* 2014; **60**(3): 189-97.

32 Azage M, Kumie A, Worku A, Bagtzoglou AC. Childhood diarrhea in high and low hotspot districts of Amhara Region, northwest Ethiopia: a multilevel modeling. *Journal of health, population, and nutrition* 2016; **35**: 13.

33 Azage M, Kumie A, Worku A, C Bagtzoglou A, Anagnostou E. Effect of climatic variability on childhood diarrhea and its high risk periods in northwestern parts of Ethiopia. *PloS one* 2017; **12**(10): e0186933.

34 Bado AR, Susuman AS, Nebie EI. Trends and risk factors for childhood diarrhea in sub-Saharan countries (1990–2013): assessing the neighborhood inequalities. *Global health action* 2016; **9**(1): 30166.

35 Banerjee B, Hazra S, B, yopadhyay D. Diarrhea management among under fives. *Indian pediatrics* 2004; **41**(3): 255-60.

36 Baqui AH, Zaman K, Persson LA, et al. Simultaneous weekly supplementation of iron and zinc is associated with lower morbidity due to diarrhea and acute lower respiratory infection in Bangladeshi infants. *The Journal of nutrition* 2003; **133**(12): 4150-7.

37 Barakat AA, Nada KH, Ezzat DA. Prevalence and determining factors of anemia and malnutrition among Egyptian children. *Indian journal of medical sciences* 2013; **67**(7): 168-77.

38 Barreto ML, Milroy CA, Strina A, et al. Community-based monitoring of diarrhea in urban Brazilian children: incidence and associated pathogens. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2006; **100**(3): 234-42.

39 Bawankule R, Singh A, Kumar K, Pedgaonkar S. Disposal of children's stools and its association with childhood diarrhea in India. *BMC public health* 2017; **17**(1): 12.

40 Becker-Dreps S, Bucardo F, Vilchez S, et al. Etiology of childhood diarrhea after rotavirus vaccine introduction: a prospective, population-based study in Nicaragua. *The Pediatric infectious disease journal* 2014; **33**(11): 1156-63.

41 Becker-Dreps S, Melendez M, Liu L, et al. Community diarrhea incidence before and after rotavirus vaccine introduction in Nicaragua. *The American journal of tropical medicine and hygiene* 2013; **89**(2): 246-50.

42 Becker-Dreps S, Paniagua M, Dominik R, et al. Changes in childhood diarrhea incidence in nicaragua following 3 years of universal infant rotavirus immunization. *The Pediatric infectious disease journal* 2011; **30**(3): 243-7.

43 Becquey E, Ouedraogo CT, Hess SY, et al. Comparison of Preventive and Therapeutic Zinc Supplementation in Young Children in Burkina Faso: A Cluster-Randomized, Community-Based Trial. *The Journal of nutrition* 2016; **146**(10): 2058-66.

44 Bennett A, Epstein LD, Gilman RH, et al. Effects of the 1997-1998 El Nino episode on community rates of diarrhea. *American journal of public health* 2012; **102**(7): e63-9.

45 Binka FN, Anto FK, Oduro AR, et al. Incidence and risk factors of paediatric rotavirus diarrhoea in northern Ghana. *Tropical medicine & international health : TM & IH* 2003; **8**(9): 840-6.

46 Biritwum RB, Asante A, Amoo PK, et al. Community-based cluster surveys on treatment preferences for diarrhoea, severe diarrhoea, and dysentery in children aged less than five years in two districts of Ghana. *Journal of health, population, and nutrition* 2004; **22**(2): 182-90.

47 Bitew BD, Woldu W, Gizaw Z. Childhood diarrheal morbidity and sanitation predictors in a nomadic community. *Italian journal of pediatrics* 2017; **43**(1): 91.

48 Bizuneh H, Getnet F, Meressa B, Tegene Y, Worku G. Factors associated with diarrheal morbidity among under-five children in Jigjiga town, Somali Regional State, eastern Ethiopia: a cross-sectional study. *BMC pediatrics* 2017; **17**(1): 182.

49 Boadi KO, Kuitunen M. Environmental and health impacts of household solid waste handling and disposal practices in third world cities: the case of the Accra Metropolitan Area, Ghana. *Journal of environmental health* 2005; **68**(4): 32-6.

50 Boisson S, Schmidt W-P, Berhanu T, Gezahegn H, Clasen T. Randomized controlled trial in rural Ethiopia to assess a portable water treatment device. *Environmental science & technology* 2009; **43**(15): 5934-9.

51 Bozkurt AI, Ozgur S, Ozcirpici B. Association between household conditions and diarrheal diseases among children in Turkey: a cohort study. *Pediatrics international : official journal of the Japan Pediatric Society* 2003; **45**(4): 443-51.

52 Brenner JL, Kabakyenga J, Kyomuhangi T, et al. Can volunteer community health workers decrease child morbidity and mortality in southwestern Uganda? An impact evaluation. *PloS one* 2011; **6**(12): e27997.

53 Briceno B, Coville A, Gertler P, Martinez S. Are there synergies from combining hygiene and sanitation promotion campaigns: Evidence from a large-scale cluster-randomized trial in rural Tanzania. *PloS one* 2017; **12**(11): e0186228.

54 Brooks WA, Santosham M, Naheed A, et al. Effect of weekly zinc supplements on incidence of pneumonia and diarrhoea in children younger than 2 years in an urban, low-income population in Bangladesh: randomised controlled trial. *Lancet (London, England)* 2005; **366**(9490): 999-1004.

55 Brown JM, Proum S, Sobsey MD. Escherichia coli in household drinking water and diarrheal disease risk: evidence from Cambodia. *Water science and technology : a*

journal of the International Association on Water Pollution Research 2008; **58**(4): 757-63.

56 Cao X, Rawalai K, Thompson AJ, et al. Relationship between feeding practices and weanling diarrhoea in northeast Thailand. *Journal of health, population, and nutrition* 2000; **18**(2): 85-92.

57 Capuno JJ, Tan CAR, Jr., Fabella VM. Do piped water and flush toilets prevent child diarrhea in rural Philippines? *Asia-Pacific journal of public health* 2015; **27**(2): NP2122-32.

58 Carlton EJ, Eisenberg JNS, Goldstick J, Cevallos W, Trostle J, Levy K. Heavy rainfall events and diarrhea incidence: the role of social and environmental factors. *American journal of epidemiology* 2014; **179**(3): 344-52.

59 Cha S, Kang D, Tuffuor B, et al. The Effect of Improved Water Supply on Diarrhea Prevalence of Children under Five in the Volta Region of Ghana: A Cluster-Randomized Controlled Trial. *International journal of environmental research and public health* 2015; **12**(10): 12127-43.

60 Cha S, Lee JE, Seo DS, et al. Associations between Household Latrines and the Prevalence of Diarrhea in Idiofa, Democratic Republic of the Congo: A Cross-Sectional Study. *The American journal of tropical medicine and hygiene* 2017; **97**(2): 460-8.

61 Checkley W, Epstein LD, Gilman RH, Cabrera L, Black RE. Effects of acute diarrhea on linear growth in Peruvian children. *American journal of epidemiology* 2003; **157**(2): 166-75.

62 Checkley W, Gilman RH, Black RE, et al. Effect of water and sanitation on childhood health in a poor Peruvian peri-urban community. *Lancet (London, England)* 2004; **363**(9403): 112-8.

63 Checkley W, Gilman RH, Black RE, et al. Effects of nutritional status on diarrhea in Peruvian children. *The Journal of pediatrics* 2002; **140**(2): 210-8.

64 Choi SYP. Mechanisms of racial inequalities in prevalence of diarrhoea in South Africa. *Journal of health, population, and nutrition* 2003; **21**(3): 264-72.

65 Chowdhury F, Khan IA, Patel S, et al. Diarrheal Illness and Healthcare Seeking Behavior among a Population at High Risk for Diarrhea in Dhaka, Bangladesh. *PloS one* 2015; **10**(6): e0130105.

66 Cifuentes E, Mazari-Hiriart M, Carneiro F, o, Bianchi F, Gonzalez D. The risk of enteric diseases in young children and environmental indicators in sentinel areas of Mexico City. *International journal of environmental health research* 2002; **12**(1): 53-62.

67 Clasen TF, Brown J, Collin S, Suntura O, Cairncross S, y. Reducing diarrhea through the use of household-based ceramic water filters: a randomized, controlled trial in rural Bolivia. *The American journal of tropical medicine and hygiene* 2004; **70**(6): 651-7.

68 Clasen T, Boisson S, Routray P, et al. Effectiveness of a rural sanitation programme on diarrhoea, soil-transmitted helminth infection, and child malnutrition in Odisha, India: a cluster-randomised trial. *The Lancet Global health* 2014; **2**(11): e645-53.

69 Coles CL, Seidman JC, Levens J, Mkocha H, Munoz B, West S. Association of mass treatment with azithromycin in trachoma-endemic communities with short-term reduced risk of diarrhea in young children. *The American journal of tropical medicine and hygiene* 2011; **85**(4): 691-6.

70 Colombara DV, H, ez B, et al. Diarrhea Prevalence, Care, and Risk Factors Among Poor Children Under 5 Years of Age in Mesoamerica. *The American journal of tropical medicine and hygiene* 2016; **94**(3): 544-52.

71 Colombara DV, Hernández B, McNellan CR, et al. Diarrhea prevalence, care, and risk factors among poor children under 5 years of age in Mesoamerica. *The American journal of tropical medicine and hygiene* 2016; **94**(3): 544-52.

72 Contreras JD, Meza R, Siebe C, et al. Health risks from exposure to untreated wastewater used for irrigation in the Mezquital Valley, Mexico: A 25-year update. *Water research* 2017; **123**: 834-50.

73 Crump JA, Mendoza CE, Priest JW, et al. Comparing serologic response against enteric pathogens with reported diarrhea to assess the impact of improved household drinking water quality. *The American journal of tropical medicine and hygiene* 2007; **77**(1): 136-41.

74 Cupul-Uicab LA, Terrazas-Medina EA, H, ez-Avila M, Longnecker MP. In utero exposure to DDT and incidence of diarrhea among boys from tropical Mexico. *Environmental research* 2017; **159**: 331-7.

75 Davis WW, Mullany LC, Shwe Oo EK, Richards AK, Iacopino V, Beyrer C. Health and Human Rights in Karen State, Eastern Myanmar. *PloS one* 2015; **10**(8): e0133822.

76 Desmennu AT, Oluwasanu MM, John-Akinola YO, Oladunni O, Adebowale AS. Maternal Education and Diarrhea among Children aged 0-24 Months in Nigeria. *African journal of reproductive health* 2017; **21**(3): 27-36.

77 Diness BR, Christoffersen D, Pedersen UB, et al. The effect of high-dose vitamin A supplementation given with bacille Calmette-Guerin vaccine at birth on infant rotavirus infection and diarrhea: a randomized prospective study from Guinea-Bissau. *The Journal of infectious diseases* 2010; **202**: S243-51.

78 Diouf K, Tabatabai P, Rudolph J, Marx M. Diarrhoea prevalence in children under five years of age in rural Burundi: an assessment of social and behavioural factors at the household level. *Global health action* 2014; **7**: 24895.

79 do Carmo GMI, Yen C, Cortes J, et al. Decline in diarrhea mortality and admissions after routine childhood rotavirus immunization in Brazil: a time-series analysis. *PLoS medicine* 2011; **8**(4): e1001024.

80 Doocy S, Burnham G. Point-of-use water treatment and diarrhoea reduction in the emergency context: an effectiveness trial in Liberia. *Tropical medicine & international health : TM & IH* 2006; **11**(10): 1542-52.

81 Dreibelbis R, Freeman MC, Greene LE, Saboori S, Rheingans R. The impact of school water, sanitation, and hygiene interventions on the health of younger siblings of pupils: a cluster-randomized trial in Kenya. *American journal of public health* 2014; **104**(1): e91-7.

82 du Preez M, Conroy RM, Ligondo S, et al. Randomized intervention study of solar disinfection of drinking water in the prevention of dysentery in Kenyan children aged under 5 years. *Environmental science & technology* 2011; **45**(21): 9315-23.

83 Pande S, Keyzer MA, Arouna A, Sonneveld BGJS. Addressing diarrhea prevalence in the West African Middle Belt: social and geographic dimensions in a case study for Benin; 2008.

84 El-Gilany AH, Hammad S. Epidemiology of diarrhoeal diseases among children under age 5 years in Dakahlia, Egypt. *Eastern Mediterranean health journal = La revue de sante de la Mediterranee orientale = al-Majallah al-sihhiyah li-sharq al-mutawassit* 2005; **11**(4): 762-75.

85 Elfadaly HA, Hassanain NA, Hassanain MA, Barakat AM, Shaapan RM. Evaluation of primitive ground water supplies as a risk factor for the development of major waterborne zoonosis in Egyptian children living in rural areas. *Journal of infection and public health* 2018; **11**(2): 203-8.

86 Elsanousi S, Abdelrahman S, Elshiekh I, et al. A study of the use and impacts of LifeStraw in a settlement camp in southern Gezira, Sudan. *Journal of water and health* 2009; **7**(3): 478-83.

87 Ercumen A, Arnold BF, Kumpel E, et al. Upgrading a piped water supply from intermittent to continuous delivery and association with waterborne illness: a matched cohort study in urban India. *PLoS medicine* 2015; **12**(10): e1001892.

88 Ercumen A, Arnold BF, Naser AM, Unicomb L, Colford JM, Jr., Luby SP. Potential sources of bias in the use of Escherichia coli to measure waterborne diarrhoea risk in low-income settings. *Tropical medicine & international health : TM & IH* 2017; **22**(1): 2-11.

89 Ercumen A, Naser AM, Unicomb L, Arnold BF, Colford JM, Jr., Luby SP. Effects of source- versus household contamination of tubewell water on child diarrhea in rural Bangladesh: a randomized controlled trial. *PloS one* 2015; **10**(3): e0121907.

90 Escamilla V, Wagner B, Yunus M, Streatfield PK, van Geen A, Emch M. Effect of deep tube well use on childhood diarrhoea in Bangladesh. *Bulletin of the World Health Organization* 2011; **89**(7): 521-7.

91 Escobar AL, Coimbra CEA, Jr., Welch JR, Horta BL, Santos RV, Cardoso AM. Diarrhea and health inequity among Indigenous children in Brazil: results from the First National Survey of Indigenous People's Health and Nutrition. *BMC public health* 2015; **15**: 191.

92 Eshete WB. A stepwise regression analysis on under-five diarrhoael morbidity prevalence in Nekemte town, western Ethiopia: maternal care giving and hygiene behavioral determinants. *East African journal of public health* 2008; **5**(3): 193-8.

93 Fabiszewski de Aceituno AM, Stauber CE, Walters AR, Meza Sanchez RE, Sobsey MD. A randomized controlled trial of the plastic-housing BioSand filter and its

impact on diarrheal disease in Copan, Honduras. *The American journal of tropical medicine and hygiene* 2012; **86**(6): 913-21.

94 Fassinou P, Elenga N, Rouet F, et al. Highly active antiretroviral therapies among HIV-1-infected children in Abidjan, Cote d'Ivoire. *AIDS (London, England)* 2004; **18**(14): 1905-13.

95 Fawzi WW, Msamanga GI, Wei R, et al. Effect of providing vitamin supplements to human immunodeficiency virus-infected, lactating mothers on the child's morbidity and CD4+ cell counts. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* 2003; **36**(8): 1053-62.

96 Fawzy A, Arpadi S, Kankasa C, et al. Early weaning increases diarrhea morbidity and mortality among uninfected children born to HIV-infected mothers in Zambia. *The Journal of infectious diseases* 2011; **203**(9): 1222-30.

97 Feleke H, Medhin G, Abebe A, Beyene B, Kloos H, Asrat D. Enteric pathogens and associated risk factors among under-five children with and without diarrhea in Wegera District, Northwestern Ethiopia. *The Pan African medical journal* 2018; **29**: 72.

98 Few R, Lake I, Hunter PR, Tran PG. Seasonality, disease and behavior: using multiple methods to explore socio-environmental health risks in the Mekong Delta. *Social science & medicine (1982)* 2013; **80**: 1-9.

99 Francis MR, Sarkar R, Roy S, et al. Effectiveness of membrane filtration to improve drinking water: A quasi-experimental study from rural southern India. *American Journal of Tropical Medicine and Hygiene* 2016; **95**(5): 1192-200.

100 Fuller JA, Clasen T, Heijnen M, Eisenberg JN. Shared sanitation and the prevalence of diarrhea in young children: evidence from 51 countries, 2001–2011. *The American journal of tropical medicine and hygiene* 2014; **91**(1): 173-80.

101 Gao W, Liu X, Yan H. Prevalence of diarrhea among children less than 36 months of age in rural western China in 2001 and 2005. *The American journal of tropical medicine and hygiene* 2014; **91**(6): 1197-202.

102 Garrett V, Ogutu P, Mabonga P, et al. Diarrhoea prevention in a high-risk rural Kenyan population through point-of-use chlorination, safe water storage, sanitation, and rainwater harvesting. *Epidemiology and infection* 2008; **136**(11): 1463-71.

103 Gebru T, Taha M, Kassahun W. Risk factors of diarrhoeal disease in under-five children among health extension model and non-model families in Sheko district rural community, Southwest Ethiopia: comparative cross-sectional study. *BMC public health* 2014; **14**: 395.

104 Genser B, Strina A, Teles CA, Prado MS, Barreto ML. Risk factors for childhood diarrhea incidence: dynamic analysis of a longitudinal study. *Epidemiology (Cambridge, Mass)* 2006; **17**(6): 658-67.

105 George CM, Perin J, Neiswender de Calani KJ, et al. Risk factors for diarrhea in children under five years of age residing in peri-urban communities in Cochabamba, Bolivia. *The American journal of tropical medicine and hygiene* 2014; **91**(6): 1190-6.

106 Geresomo NC, Mbuthia EK, Matofari JW, Mwangwela AM. Targeting caregivers with context specific behavior change training increased uptake of recommended hygiene practices during food preparation and complementary feeding in Dedza district of Central Malawi. *Ecology of food and nutrition* 2018; **57**(4): 301-13.

107 Ghimire M, Pradhan YV, Maskey MK. Community-based interventions for diarrhoeal diseases and acute respiratory infections in Nepal. *Bulletin of the World Health Organization* 2010; **88**(3): 216-21.

108 Gorham TJ, Yoo J, Garabed R, Mouhaman A, Lee J. Water Access, Sanitation, and Hygiene Conditions and Health Outcomes among Two Settlement Types in Rural Far North Cameroon. *International journal of environmental research and public health* 2017; **14**(4).

109 Graf J, Meierhofer R, Wegelin M, Mosler H-J. Water disinfection and hygiene behaviour in an urban slum in Kenya: impact on childhood diarrhoea and influence of beliefs. *International journal of environmental health research* 2008; **18**(5): 335-55.

110 Graf J, Zebaze Togouet S, Kemka N, Niyitegeka D, Meierhofer R, Gangoue Pieboji J. Health gains from solar water disinfection (SODIS): evaluation of a water quality intervention in Yaounde, Cameroon. *Journal of water and health* 2010; **8**(4): 779-96.

111 Gundry SW, Wright JA, Conroy RM, et al. Child dysentery in the Limpopo Valley: a cohort study of water, sanitation and hygiene risk factors. *Journal of water and health* 2009; **7**(2): 259-66.

112 Guo Y-F, Gan Y-Y, Guo C-N, Sun J, Hao L-P. Nutritional status of under-five children from urban low-income families in Xiangtan and Jilin in China. *Journal of Huazhong University of Science and Technology Medical sciences = Hua zhong ke ji da xue xue bao Yi xue Ying De wen ban = Huazhong keji daxue xuebao Yixue Yingdewen ban* 2017; **37**(1): 74-8.

113 Gupta DN, Mondal SK, Ghosh S, Rajendran K, Sur D, Manna B. Impact of zinc supplementation on diarrhoeal morbidity in rural children of West Bengal, India. *Acta paediatrica (Oslo, Norway : 1992)* 2003; **92**(5): 531-6.

114 Habib MA, Soofi S, Sadiq K, et al. A study to evaluate the acceptability, feasibility and impact of packaged interventions ("Diarrhea Pack") for prevention and treatment of childhood diarrhea in rural Pakistan. *BMC public health* 2013; **13**: 922.

115 Halder AK, Luby SP, Akhter S, Ghosh PK, Johnston RB, Unicomb L. Incidences and Costs of Illness for Diarrhea and Acute Respiratory Infections for Children < 5 Years of Age in Rural Bangladesh. *The American journal of tropical medicine and hygiene* 2017; **96**(4): 953-60.

116 Haque R, Mondal D, Kirkpatrick BD, et al. Epidemiologic and clinical characteristics of acute diarrhea with emphasis on Entamoeba histolytica infections in preschool children in an urban slum of Dhaka, Bangladesh. *The American journal of tropical medicine and hygiene* 2003; **69**(4): 398-405.

117 Harding E, Beckworth C, Fesselet J-F, Lenglet A, Lako R, Valadez JJ. Using lot quality assurance sampling to assess access to water, sanitation and hygiene services in a refugee camp setting in South Sudan: a feasibility study. *BMC public health* 2017; **17**(1): 643.

118 Harshfield E, Lantagne D, Turbes A, Null C. Evaluating the sustained health impact of household chlorination of drinking water in rural Haiti. *The American journal of tropical medicine and hygiene* 2012; **87**(5): 786-95.

119 Hartinger SM, Lanata CF, Hattendorf J, et al. Improving household air, drinking water and hygiene in rural Peru: a community-randomized-controlled trial of an integrated environmental home-based intervention package to improve child health. *International journal of epidemiology* 2016; **45**(6): 2089-99.

120 Hasan MM, Richardson A. How sustainable household environment and knowledge of healthy practices relate to childhood morbidity in South Asia: analysis of survey data from Bangladesh, Nepal and Pakistan. *BMJ open* 2017; **7**(6): e015019.

121 Hassan KE, Mansour A, Shaheen H, et al. The impact of household hygiene on the risk of bacterial diarrhea among Egyptian children in rural areas, 2004-2007. *Journal of infection in developing countries* 2014; **8**(12): 1541-51.

122 Hatt LE, Waters HR. Determinants of child morbidity in Latin America: a pooled analysis of interactions between parental education and economic status. *Social science & medicine (1982)* 2006; **62**(2): 375-86.

123 Headey D, Nguyen P, Kim S, Rawat R, Ruel M, Menon P. Is Exposure to Animal Feces Harmful to Child Nutrition and Health Outcomes? A Multicountry Observational Analysis. *The American journal of tropical medicine and hygiene* 2017; **96**(4): 961-9.

124 Hess SY, Abbeddou S, Jimenez EY, et al. Small-quantity lipid-based nutrient supplements, regardless of their zinc content, increase growth and reduce the prevalence of stunting and wasting in young burkinabe children: a cluster-randomized trial. *PloS one* 2015; **10**(3): e0122242.

125 Hipgrave DB, Assefa F, Winoto A, Sukotjo S. Donated breast milk substitutes and incidence of diarrhoea among infants and young children after the May 2006 earthquake in Yogyakarta and Central Java. *Public health nutrition* 2012; **15**(2): 307-15.

126 Hollm-Delgado M-G, Gilman RH, Bern C, et al. Lack of an adverse effect of Giardia intestinalis infection on the health of Peruvian children. *American journal of epidemiology* 2008; **168**(6): 647-55.

127 Hong TK, Dibley MJ, Tuan T. Factors affecting utilization of health care services by mothers of children ill with diarrhea in rural Vietnam. *The Southeast Asian journal of tropical medicine and public health* 2003; **34**(1): 187-98.

128 Houatthongkham S, Sithivong N, Jennings G, et al. Trends in the incidence of acute watery diarrhoea in the Lao People's Democratic Republic, 2009-2013. *Western Pacific surveillance and response journal : WPSAR* 2016; **7**(3): 6-14.

129 Huda TMN, Unicomb L, Johnston RB, Halder AK, Yushuf Sharker MA, Luby SP. Interim evaluation of a large scale sanitation, hygiene and water improvement

programme on childhood diarrhea and respiratory disease in rural Bangladesh. *Social science & medicine (1982)* 2012; **75**(4): 604-11.

130 Hunter PR, Risebro H, Yen M, et al. Water source and diarrhoeal disease risk in children under 5 years old in Cambodia: a prospective diary based study. *BMC public health* 2013; **13**: 1145.

131 Iannotti LL, Zavaleta N, Leon Z, Caulfield LE. Growth and body composition of Peruvian infants in a periurban setting. *Food and nutrition bulletin* 2009; **30**(3): 245-53.

132 Imada KS, Araujo TSd, Muniz PT, Padua VLd. Socioeconomic, hygienic, and sanitation factors in reducing diarrhea in the Amazon. *Revista de saude publica* 2016; **50**: 77.

133 Isenbarger DW, Hien BT, Ha HT, et al. Prospective study of the incidence of diarrhoea and prevalence of bacterial pathogens in a cohort of Vietnamese children along the Red River. *Epidemiology and infection* 2001; **127**(2): 229-36.

134 Islam M, Ercumen A, Ashraf S, et al. Unsafe disposal of feces of children <3 years among households with latrine access in rural Bangladesh: Association with household characteristics, fly presence and child diarrhea. *PloS one* 2018; **13**(4): e0195218.

135 Islam M, Ercumen A, Naser AM, et al. Effectiveness of the Hydrogen Sulfide Test as a Water Quality Indicator for Diarrhea Risk in Rural Bangladesh. *The American journal of tropical medicine and hygiene* 2017; **97**(6): 1867-71.

136 Jaeggi T, Kortman GAM, Moretti D, et al. Iron fortification adversely affects the gut microbiome, increases pathogen abundance and induces intestinal inflammation in Kenyan infants. *Gut* 2015; **64**(5): 731-42.

137 Jensen PK, Ensink JHJ, Jayasinghe G, et al. Effect of chlorination of drinking-water on water quality and childhood diarrhoea in a village in Pakistan. *Journal of health, population, and nutrition* 2003; **21**(1): 26-31.

138 Kaindi DWM, Schelling E, Wangoh JM, Imungi JK, Farah Z, Meile L. Risk factors for symptoms of gastrointestinal illness in rural town Isiolo, Kenya. *Zoonoses and public health* 2012; **59**(2): 118-25.

139 Kamara JK, Galuk, e M, Maeda F, Luboga S, Renzaho A. Understanding the challenges of improving sanitation and hygiene outcomes in a community based intervention: A cross-sectional study in rural Tanzania. *International journal of environmental research and public health* 2017; **14**(6): 602.

140 Kamm KB, Feikin DR, Bigogo GM, et al. Associations between presence of handwashing stations and soap in the home and diarrhoea and respiratory illness, in children less than five years old in rural western Kenya. *Tropical medicine & international health : TM & IH* 2014; **19**(4): 398-406.

141 Kante AM, Gutierrez HR, Larsen AM, et al. Childhood Illness Prevalence and Health Seeking Behavior Patterns in Rural Tanzania. *BMC public health* 2015; **15**: 951.

142 Kanté AM, Gutierrez HR, Larsen AM, et al. Childhood illness prevalence and health seeking behavior patterns in rural Tanzania. *BMC public health* 2015; **15**(1): 951.

143 Kariuki JG, Magambo KJ, Njeruh MF, Muchiri EM, Nzioka SM, Kariuki S. Effects of hygiene and sanitation interventions on reducing diarrhoea prevalence among children in resource constrained communities: case study of Turkana District, Kenya. *Journal of community health* 2012; **37**(6): 1178-84.

144 Kattula D, Francis MR, Kulinkina A, et al. Environmental predictors of diarrhoeal infection for rural and urban communities in south India in children and adults. *Epidemiology and infection* 2015; **143**(14): 3036-47.

145 Kawakatsu Y, Tanaka J, Ogawa K, Ogendo K, Honda S. Community unit performance: factors associated with childhood diarrhea and appropriate treatment in Nyanza Province, Kenya. *BMC public health* 2017; **17**(1): 202.

146 Khan MN, Islam MM. Effect of exclusive breastfeeding on selected adverse health and nutritional outcomes: a nationally representative study. *BMC public health* 2017; **17**(1): 889.

147 Kirby MA, Nagel CL, Rosa G, et al. Use, microbiological effectiveness and health impact of a household water filter intervention in rural Rwanda-A matched cohort study. *International journal of hygiene and environmental health* 2017; **220**(6): 1020-9.

148 Kolahi A-A, Rastegarpour A, Abadi A, Gachkar L. An unexpectedly high incidence of acute childhood diarrhea in Koot-Abdollah, Ahwaz, Iran. *International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases* 2010; **14**(7): e618-21.

149 Kolahi A-A, Rastegarpour A, Sohrabi M-R. The impact of an urban sewerage system on childhood diarrhoea in Tehran, Iran: a concurrent control field trial. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2009; **103**(5): 500-5.

150 Komarulzaman A, Smits J, de Jong E. Clean water, sanitation and diarrhoea in Indonesia: Effects of household and community factors. *Global public health* 2017; **12**(9): 1141-55.

151 Kosek M, Yori PP, Pan WK, et al. Epidemiology of highly endemic multiply antibiotic-resistant shigellosis in children in the Peruvian Amazon. *Pediatrics* 2008; **122**(3): e541-9.

152 Kulinkina A, V. r, Mohan VR, et al. Seasonality of water quality and diarrheal disease counts in urban and rural settings in south India. *Scientific reports* 2016; **6**: 20521.

153 Kumar S, Vollmer S. Does access to improved sanitation reduce childhood diarrhea in rural India? *Health economics* 2013; **22**(4): 410-27.

154 Kumi-Kyereme A, Amo-Adjei J. Household wealth, residential status and the incidence of diarrhoea among children under-five years in Ghana. *Journal of epidemiology and global health* 2016; **6**(3): 131-40.

155 Lee G, Cama V, Gilman RH, et al. Comparison of two types of epidemiological surveys aimed at collecting daily clinical symptoms in community-based longitudinal studies. *Annals of epidemiology* 2010; **20**(2): 151-8.

156 Lee G, Yori P, Olortegui MP, et al. Comparative effects of vivax malaria, fever and diarrhoea on child growth. *International journal of epidemiology* 2012; **41**(2): 531-9.

157 Lee H-Y, Van Huy N, Choi S. Determinants of early childhood morbidity and proper treatment responses in Vietnam: results from the Multiple Indicator Cluster Surveys, 2000-2011. *Global health action* 2016; **9**: 29304.

158 Lindquist ED, George CM, Perin J, et al. A cluster randomized controlled trial to reduce childhood diarrhea using hollow fiber water filter and/or hygiene-sanitation educational interventions. *The American journal of tropical medicine and hygiene* 2014; **91**(1): 190-7.

159 Long KZ, Montoya Y, Hertzmark E, Santos JI, Rosado JL. A double-blind, randomized, clinical trial of the effect of vitamin A and zinc supplementation on diarrheal disease and respiratory tract infections in children in Mexico City, Mexico. *The American journal of clinical nutrition* 2006; **83**(3): 693-700.

160 Long KZ, Rosado JL, DuPont HL, Hertzmark E, Santos JI. Supplementation with vitamin A reduces watery diarrhoea and respiratory infections in Mexican children. *The British journal of nutrition* 2007; **97**(2): 337-43.

161 Luabeya K-KA, Mpontshane N, Mackay M, et al. Zinc or multiple micronutrient supplementation to reduce diarrhea and respiratory disease in South African children: a randomized controlled trial. *PloS one* 2007; **2**(6): e541.

162 Luby SP, Agboatwalla M, Painter J, Altaf A, Billhimer WL, Hoekstra RM. Effect of intensive handwashing promotion on childhood diarrhea in high-risk communities in Pakistan: a randomized controlled trial. *JAMA* 2004; **291**(21): 2547-54.

163 Luby SP, Halder AK, Huda TM, et al. Microbiological Contamination of Drinking Water Associated with Subsequent Child Diarrhea. *The American journal of tropical medicine and hygiene* 2015; **93**(5): 904-11.

164 Luby SP, Rahman M, Arnold BF, et al. Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Bangladesh: a cluster randomised controlled trial. *The Lancet Global health* 2018; **6**(3): e302-e15.

165 Mach O, Lu L, Creek T, et al. Population-based study of a widespread outbreak of diarrhea associated with increased mortality and malnutrition in Botswana, January-March, 2006. *The American journal of tropical medicine and hygiene* 2009; **80**(5): 812-8.

166 Mahmud MA, Hossain MM, Huang DB, Habib M, DuPont HL. Sociodemographic, environmental and clinical risk factors for developing persistent diarrhoea among infants in a rural community of Egypt. *Journal of health, population, and nutrition* 2001; **19**(4): 313-9.

167 Malik A, Taneja DK, Devasenapathy N, Rajeshwari K. Short-course prophylactic zinc supplementation for diarrhea morbidity in infants of 6 to 11 months. *Pediatrics* 2013; **132**(1): e46-52.

168 Manger MS, Taneja S, S, et al. Poor folate status predicts persistent diarrhea in 6- to 30-month-old north Indian children. *The Journal of nutrition* 2011; **141**(12): 2226-32.

169 Manna B, Nasrin D, Kanungo S, et al. Determinants of health care seeking for diarrheal illness in young children in urban slums of Kolkata, India. *The American journal of tropical medicine and hygiene* 2013; **89**(1): 56-61.

170 Mansour A, Shaheen HI, Amine M, et al. Diarrhea burden due to natural infection with enterotoxigenic Escherichia coli in a birth cohort in a rural Egyptian community. *Journal of clinical microbiology* 2014; **52**(7): 2595-603.

171 Marcynuk PB, Flint JA, Sargeant JM, et al. Comparison of the burden of diarrhoeal illness among individuals with and without household cisterns in northeast Brazil. *BMC infectious diseases* 2013; **13**: 65.

172 Markovitz A, R. a, Goldstick JE, et al. Where science meets policy: comparing longitudinal and cross-sectional designs to address diarrhoeal disease burden in the developing world. *International journal of epidemiology* 2012; **41**(2): 504-13.

173 Masangwi SJ, Ferguson NS, Grimason AM, Morse TD, Zawdie G, Kazembe LN. Household and community variations and nested risk factors for diarrhoea prevalence in southern Malawi: a binary logistic multi-level analysis. *International journal of environmental health research* 2010; **20**(2): 141-58.

174 Mashal T, Takano T, Nakamura K, et al. Factors associated with the health and nutritional status of children under 5 years of age in Afghanistan: family behaviour related to women and past experience of war-related hardships. *BMC public health* 2008; **8**: 301.

175 Mashoto KO, Malebo HM, Msisiri E, Peter E. Prevalence, one week incidence and knowledge on causes of diarrhea: household survey of under-fives and adults in Mkuranga district, Tanzania. *BMC public health* 2014; **14**: 985.

176 Mazumder S, Taneja S, Bahl R, et al. Effect of implementation of integrated management of neonatal and childhood illness programme on treatment seeking practices for morbidities in infants: cluster randomised trial. *BMJ (Clinical research ed)* 2014; **349**: g4988.

177 Mbonye AK. Prevalence of childhood illnesses and care-seeking practices in rural Uganda. *TheScientificWorldJournal* 2003; **3**: 721-30.

178 Mbonye AK. Risk factors for diarrhoea and upper respiratory tract infections among children in a rural area of Uganda. *Journal of health, population, and nutrition* 2004; **22**(1): 52-8.

179 Mekasha A, Tesfahun A. Determinants of diarrhoeal diseases: a community based study in urban south western Ethiopia. *East African medical journal* 2003; **80**(2): 77-82.

180 Melo MCNd, Taddei JAAC, Diniz-Santos DR, et al. Incidence of diarrhea in children living in urban slums in Salvador, Brazil. *The Brazilian journal of infectious diseases : an official publication of the Brazilian Society of Infectious Diseases* 2008; **12**(1): 89-93.

181 Melo MCNd, Taddei JAdAC, Diniz-Santos DR, May DS, Carneiro NB, Silva LR. Incidence of diarrhea: poor parental recall ability. *The Brazilian journal of infectious diseases : an official publication of the Brazilian Society of Infectious Diseases* 2007; **11**(6): 571-9.

182 Mengistie B, Berhane Y, Worku A. Household water chlorination reduces incidence of diarrhea among under-five children in rural Ethiopia: a cluster randomized controlled trial. *PloS one* 2013; **8**(10): e77887.

183 Merga N, Alemayehu T. Knowledge, perception, and management skills of mothers with under-five children about diarrhoeal disease in indigenous and resettlement communities in Assosa District, Western Ethiopia. *Journal of health, population, and nutrition* 2015; **33**(1): 20-30.

184 Mihrshahi S, Ichikawa N, Shuaib M, et al. Prevalence of exclusive breastfeeding in Bangladesh and its association with diarrhoea and acute respiratory infection: results of the multiple indicator cluster survey 2003. *Journal of health, population, and nutrition* 2007; **25**(2): 195-204.

185 Moll DM, McElroy RH, Sabogal R, Corrales LF, Gelting RJ. Health impact of water and sanitation infrastructure reconstruction programmes in eight Central American communities affected by Hurricane Mitch. *Journal of water and health* 2007; **5**(1): 51-65.

186 Molla NA, Ali G, Mollah KA, et al. Migration, health, and socioenvironmental safety net among children of Dhaka, Bangladesh. *Archives of environmental & occupational health* 2017; **72**(6): 336-42.

187 Mondal D, Haque R, Sack RB, Kirkpatrick BD, Petri WA, Jr. Attribution of malnutrition to cause-specific diarrheal illness: evidence from a prospective study of preschool children in Mirpur, Dhaka, Bangladesh. *The American journal of tropical medicine and hygiene* 2009; **80**(5): 824-6.

188 Moore SR, Lima AA, Schorling JB, Barboza MS, Jr., Soares AM, Guerrant RL. Changes over time in the epidemiology of diarrhea and malnutrition among children in an urban Brazilian shantytown, 1989 to 1996. *International journal of infectious*

diseases : IJID : official publication of the International Society for Infectious Diseases 2000; **4**(4): 179-86.

189 Moraes LR, Cancio JA, Cairncross S, y, Huttly S. Impact of drainage and sewerage on diarrhoea in poor urban areas in Salvador, Brazil. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2003; **97**(2): 153-8.

190 Morita T, Perin J, Oldja L, et al. Mouthing of Soil Contaminated Objects is Associated with Environmental Enteropathy in Young Children. *Tropical medicine & international health : TM & IH* 2017; **22**(6): 670-8.

191 Mozumder AB, Barkat EK, Kane TT, Levin A, Ahmed S. The effect of birth interval on malnutrition in Bangladeshi infants and young children. *Journal of biosocial science* 2000; **32**(3): 289-300.

192 Mshida HA, Kassim N, Kimanya ME, Mpolya E. Influence of Water, Sanitation, and Hygiene Practices on Common Infections among Under-Five Children in Longido and Monduli Districts of Arusha, Tanzania. *Journal of environmental and public health* 2017; **2017**: 9235168.

193 Musengimana G, Mukinda FK, Machekano R, Mahomed H. Temperature Variability and Occurrence of Diarrhoea in Children under Five-Years-Old in Cape Town Metropolitan Sub-Districts. *International journal of environmental research and public health* 2016; **13**(9).

194 Nakahara S, Poudel KC, Lopchan M, et al. Differential effects of out-of-home day care in improving child nutrition and augmenting maternal income among those with and without childcare support: a prospective before-after comparison study in Pokhara, Nepal. *Health policy (Amsterdam, Netherlands)* 2010; **97**(1): 16-25.

195 Nakamori M, Nguyen XN, Nguyen CK, et al. Nutritional status, feeding practice and incidence of infectious diseases among children aged 6 to 18 months in northern mountainous Vietnam. *The journal of medical investigation : JMI* 2010; **57**(1): 45-53.

196 Ndeezi G, Tylleskar T, Ndugwa CM, Tumwine JK. Multiple micronutrient supplementation does not reduce diarrhoea morbidity in Ugandan HIV-infected children: a randomised controlled trial. *Paediatrics and international child health* 2012; **32**(1): 14-21.

197 Nikiema L, Huybregts L, Martin-Prevel Y, et al. Effectiveness of facility-based personalized maternal nutrition counseling in improving child growth and morbidity up to 18 months: A cluster-randomized controlled trial in rural Burkina Faso. *PloS one* 2017; **12**(5): e0177839.

198 Null C, Stewart CP, Pickering AJ, et al. Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Kenya: a cluster-randomised controlled trial. *The Lancet Global Health* 2018; **6**(3): e316-e29.198.

199 Oberhelman RA, Gilman RH, Sheen P, et al. An intervention-control study of corralling of free-ranging chickens to control Campylobacter infections among children in a Peruvian periurban shantytown. *The American journal of tropical medicine and hygiene* 2006; **74**(6): 1054-9.

200 Okronipa HET, Marquis GS, Lartey A, Brakohiapa L, Perez-Escamilla R, Mazur RE. Postnatal depression symptoms are associated with increased diarrhea among infants of HIV-positive Ghanaian mothers. *AIDS and behavior* 2012; **16**(8): 2216-25.

201 Olney DK, Pedehombga A, Ruel MT, Dillon A. A 2-year integrated agriculture and nutrition and health behavior change communication program targeted to women in Burkina Faso reduces anemia, wasting, and diarrhea in children 3-12.9 months of

age at baseline: a cluster-randomized controlled trial. *The Journal of nutrition* 2015; **145**(6): 1317-24.

202 Omore R, O'Reilly CE, Williamson J, et al. Health care-seeking behavior during childhood diarrheal illness: results of health care utilization and attitudes surveys of caretakers in western Kenya, 2007-2010. *The American journal of tropical medicine and hygiene* 2013; **89**(1): 29-40.

203 Opryszko MC, Majeed SW, Hansen PM, et al. Water and hygiene interventions to reduce diarrhoea in rural Afghanistan: a randomized controlled study. *Journal of water and health* 2010; **8**(4): 687-702.

204 Ortiz J, Van Camp J, Wijaya S, Donoso S, Huybregts L. Determinants of child malnutrition in rural and urban Ecuadorian highlands. *Public health nutrition* 2014; **17**(9): 2122-30.

205 Osendarp SJ, van Raaij JM, Darmstadt GL, Baqui AH, Hautvast JG, Fuchs GJ. Zinc supplementation during pregnancy and effects on growth and morbidity in low birthweight infants: a randomised placebo controlled trial. *Lancet (London, England)* 2001; **357**(9262): 1080-5.

206 Owusu-Agyei S, Newton S, Mahama E, et al. Impact of vitamin A with zinc supplementation on malaria morbidity in Ghana. *Nutrition journal* 2013; **12**: 131.

207 Page A-L, Hustache S, Luquero FJ, Djibo A, Manzo ML, Grais RF. Health care seeking behavior for diarrhea in children under 5 in rural Niger: results of a cross-sectional survey. *BMC public health* 2011; **11**: 389.

208 Pathela P, Zahid Hasan K, Roy E, Huq F, Kasem Siddique A, Bradley Sack R. Diarrheal illness in a cohort of children 0-2 years of age in rural Bangladesh: I. Incidence and risk factors. *Acta paediatrica (Oslo, Norway : 1992)* 2006; **95**(4): 430-7.

209 Peletz R, S, i M, et al. Drinking water quality, feeding practices, and diarrhea among children under 2 years of HIV-positive mothers in peri-urban Zambia. *The American journal of tropical medicine and hygiene* 2011; **85**(2): 318-26.

210 Pickering AJ, Djebbari H, Lopez C, Coulibaly M, Alzua ML. Effect of a community-led sanitation intervention on child diarrhoea and child growth in rural Mali: a cluster-randomised controlled trial. *The Lancet Global health* 2015; **3**(11): e701-11.

211 Piechulek H, Al-Sabbir A, Mendoza-Aldana J. Diarrhea and ARI in rural areas of Bangladesh. *The Southeast Asian journal of tropical medicine and public health* 2003; **34**(2): 337-42.

212 Plate DK, Strassmann BI, Wilson ML. Water sources are associated with childhood diarrhoea prevalence in rural east-central Mali. *Tropical medicine & international health : TM & IH* 2004; **9**(3): 416-25.

213 Profitos JMH, Mouhaman A, Lee S, et al. Muddying the waters: A new area of concern for drinking water contamination in Cameroon. *International Journal of Environmental Research and Public Health* 2014; **11**(12): 12454-72.

214 Prohmmo A, Cook LA, Murdoch DR. Childhood diarrhoea in a district in northeast Thailand: incidence and treatment choices. *Asia-Pacific journal of public health* 2006; **18**(2): 26-32.

215 Qadri F, Saha A, Ahmed T, Al Tarique A, Begum YA, Svennerholm A-M. Disease burden due to enterotoxigenic Escherichia coli in the first 2 years of life in an urban community in Bangladesh. *Infection and immunity* 2007; **75**(8): 3961-8.

216 Quadri F, Nasrin D, Khan A, et al. Health care use patterns for diarrhea in children in low-income periurban communities of Karachi, Pakistan. *The American journal of tropical medicine and hygiene* 2013; **89**(1): 49-55.

217 Rahman MM, Vermund SH, Wahed MA, Fuchs GJ, Baqui AH, Alvarez JO. Simultaneous zinc and vitamin A supplementation in Bangladeshi children: randomised double blind controlled trial. *BMJ (Clinical research ed)* 2001; **323**(7308): 314-8.

218 Rao MR, Abu-Elyazeed R, Savarino SJ, et al. High disease burden of diarrhea due to enterotoxigenic Escherichia coli among rural Egyptian infants and young children. *Journal of clinical microbiology* 2003; **41**(10): 4862-4.

219 Regassa W, Lemma S. Assessment of Diarrheal Disease Prevalence and Associated Risk Factors in Children of 6-59 Months Old at Adama District Rural Kebeles, Eastern Ethiopia, January/2015. *Ethiopian journal of health sciences* 2016; **26**(6): 581-8.

220 Rego RF, Moraes LRS, Dourado I. Diarrhoea and garbage disposal in Salvador, Brazil. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2005; **99**(1): 48-54.

221 Reller ME, Mendoza CE, Lopez MB, et al. A randomized controlled trial of household-based flocculant-disinfectant drinking water treatment for diarrhea prevention in rural Guatemala. *The American journal of tropical medicine and hygiene* 2003; **69**(4): 411-9.

222 Roberts L, Chartier Y, Chartier O, Malenga G, Toole M, Rodka H. Keeping clean water clean in a Malawi refugee camp: a randomized intervention trial. *Bulletin of the World Health Organization* 2001; **79**(4): 280-7.

223 Rogawski ET, Westreich D, Becker-Dreps S, et al. The effect of early life antibiotic exposures on diarrheal rates among young children in Vellore, India. *The Pediatric infectious disease journal* 2015; **34**(6): 583-8.

224 Rollins NC, Ndirangu J, B, et al. Exclusive breastfeeding, diarrhoeal morbidity and all-cause mortality in infants of HIV-infected and HIV uninfected mothers: an

intervention cohort study in KwaZulu Natal, South Africa. *PloS one* 2013; **8**(12): e81307.

225 Ross J, Hanlon C, Medhin G, et al. Perinatal mental distress and infant morbidity in Ethiopia: a cohort study. *Archives of disease in childhood Fetal and neonatal edition* 2011; **96**(1): F59-64.

226 Roushdy R, Sieverding M. Improved water and child health in Egypt: impact of interrupted water supply and storage of household water on the prevalence of diarrhoea. *Eastern Mediterranean health journal = La revue de sante de la Mediterranee orientale = al-Majallah al-sihhiyah li-sharq al-mutawassit* 2016; **22**(1): 4-18.

227 Saha D, Akinsola A, Sharples K, et al. Health Care Utilization and Attitudes Survey: understanding diarrheal disease in rural Gambia. *The American journal of tropical medicine and hygiene* 2013; **89**(1): 13-20.

228 Saha S. Design and baseline findings of a multi-site non-randomized evaluation of the effect of a health programme on microfinance clients in India. *Global journal of health science* 2013; **6**(1): 43-51.

229 Santos CB, Araujo KCGM, Jardim-Botelho A, et al. Diarrhea incidence and intestinal infections among rotavirus vaccinated infants from a poor area in Brazil: a spatial analysis. *BMC public health* 2014; **14**: 399.

230 Sarkar R, Gladstone BP, Warier JP, et al. Rotavirus and other Diarrheal Disease in a Birth Cohort from Southern Indian Community. *Indian pediatrics* 2016; **53**(7): 583-8.

231 Sastry N, Burgard S. The prevalence of diarrheal disease among Brazilian children: trends and differentials from 1986 to 1996. *Social science & medicine (1982)* 2005; **60**(5): 923-35.

232 Sazawal S, Dhingra U, Dhingra P, et al. Effects of fortified milk on morbidity in young children in north India: community based, randomised, double masked placebo controlled trial. *BMJ (Clinical research ed)* 2007; **334**(7585): 140.

233 Sazawal S, Dhingra U, Hiremath G, et al. Prebiotic and probiotic fortified milk in prevention of morbidities among children: community-based, randomized, double-blind, controlled trial. *PloS one* 2010; **5**(8): e12164.

234 Schmidt WP, Boisson S, Routray P, et al. Exposure to cows is not associated with diarrhoea or impaired child growth in rural Odisha, India: a cohort study. *Epidemiology and infection* 2016; **144**(1): 53-63.

235 Senarath U, Dibley MJ, Agho KE. Breast-feeding Performance Index: a composite index to describe overall breast-feeding performance among infants under 6 months of age. *Public health nutrition* 2007; **10**(10): 996-1004.

236 Shah SM, Yousafzai M, Lakhani NB, Chotani RA, Nowshad G. Prevalence and correlates of diarrhea. *Indian journal of pediatrics* 2003; **70**(3): 207-11.

237 Sheth M, Obrah M. Diarrhea prevention through food safety education. *Indian journal of pediatrics* 2004; **71**(10): 879-82.

238 Shivoga WA, Moturi WN. Geophagia as a risk factor for diarrhoea. *Journal of infection in developing countries* 2009; **3**(2): 94-8.

239 Sima LC, Desai MM, McCarty KM, Elimelech M. Relationship between use of water from community-scale water treatment refill kiosks and childhood diarrhea in Jakarta. *The American journal of tropical medicine and hygiene* 2012; **87**(6): 979-84.

240 Simanjuntak CH, Punjabi NH, Wangsasaputra F, et al. Diarrhoea episodes and treatment-seeking behaviour in a slum area of North Jakarta, Indonesia. *Journal of health, population, and nutrition* 2004; **22**(2): 119-29.

241 Soofi S, Cousens S, Iqbal SP, et al. Effect of provision of daily zinc and iron with several micronutrients on growth and morbidity among young children in Pakistan: a cluster-randomised trial. *Lancet (London, England)* 2013; **382**(9886): 29-40.

242 Sur D, Saha DR, Manna B, Rajendran K, Bhattacharya SK. Periodic deworming with albendazole and its impact on growth status and diarrhoeal incidence among children in an urban slum of India. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2005; **99**(4): 261-7.

243 Sur D, Gupta DN, Mondal SK, et al. Impact of zinc supplementation on diarrheal morbidity and growth pattern of low birth weight infants in kolkata, India: a randomized, double-blind, placebo-controlled, community-based study. *Pediatrics* 2003; **112**(6): 1327-32.

244 Sur D, Manna B, Deb AK, et al. Factors associated with reported diarrhoea episodes and treatment-seeking in an urban slum of Kolkata, India. *Journal of health, population, and nutrition* 2004; **22**(2): 130-8.

245 Takanashi K, Chonan Y, Quyen DT, Khan NC, Poudel KC, Jimba M. Survey of food-hygiene practices at home and childhood diarrhoea in Hanoi, Viet Nam. *Journal of health, population, and nutrition* 2009; **27**(5): 602-11.

246 Takanashi K, Quyen DT, Le Hoa NT, Khan NC, Yasuoka J, Jimba M. Long-term impact of community-based information, education and communication activities on food hygiene and food safety behaviors in Vietnam: a longitudinal study. *PloS one* 2013; **8**(8): e70654.

247 Talal AH, Moe CL, Lima AA, et al. Seroprevalence and seroincidence of Norwalk-like virus infection among Brazilian infants and children. *Journal of medical virology* 2000; **61**(1): 117-24.

248 Taneja S, B, ari N, et al. Effect of zinc supplementation on morbidity and growth in hospital-born, low-birth-weight infants. *The American journal of clinical nutrition* 2009; **90**(2): 385-91.

249 Taneja S, S, TA, et al. Folic acid and vitamin B-12 supplementation and common infections in 6-30-mo-old children in India: a randomized placebo-controlled trial. *The American journal of clinical nutrition* 2013; **98**(3): 731-7.

250 Tawfeek HI, Najim NH, Al-Mashikhi S. Efficacy of an infant formula containing anti-Escherichia coli colostral antibodies from hyperimmunized cows in preventing diarrhea in infants and children: a field trial. *International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases* 2003; **7**(2): 120-8.

251 Thiam S, Diene AN, Fuhrimann S, et al. Prevalence of diarrhoea and risk factors among children under five years old in Mbour, Senegal: a cross-sectional study. *Infectious diseases of poverty* 2017; **6**(1): 109.

252 Thiem VD, Schmidt W-P, Suzuki M, et al. Animal livestock and the risk of hospitalized diarrhoea in children under 5 years in Vietnam. *Tropical medicine & international health : TM & IH* 2012; **17**(5): 613-21.

253 Tornheim JA, M, KB, L, rigan PJ, Cifuentes E. Water privatization, water source, and pediatric diarrhea in Bolivia: epidemiologic analysis of a social experiment. *International journal of occupational and environmental health* 2009; **15**(3): 241-8.

254 Trudeau J, Aksan A-M, Vasquez WF. Water system unreliability and diarrhea incidence among children in Guatemala. *International journal of public health* 2018; **63**(2): 241-50.

255 Tsiko RG. Bayesian spatial analysis of childhood diseases in Zimbabwe. *BMC public health* 2015; **15**: 842.

256 Turab A, Soofi SB, Ahmed I, Bhatti Z, Zaidi AK, Bhutta ZA. Demographic, socioeconomic, and health characteristics of the MAL-ED network study site in rural Pakistan. *Clinical Infectious Diseases* 2014; **59**: S304-S9.

257 Tylleskar T, Jackson D, Meda N, et al. Exclusive breastfeeding promotion by peer counsellors in sub-Saharan Africa (PROMISE-EBF): a cluster-randomised trial. *Lancet (London, England)* 2011; **378**(9789): 420-7.

258 van der Hoek W, Konradsen F, Ensink JH, Mudasser M, Jensen PK. Irrigation water as a source of drinking water: is safe use possible? *Tropical medicine & international health : TM & IH* 2001; **6**(1): 46-54.

259 van der Hoek W, Feenstra SG, Konradsen F. Availability of irrigation water for domestic use in Pakistan: its impact on prevalence of diarrhoea and nutritional status of children. *Journal of health, population, and nutrition* 2002; **20**(1): 77-84.

260 Vasconcelos MJdOB, Rissin A, Figueiroa JN, Lira PICd, Batista Filho M. Factors associated with diarrhea in children under five years old in the state of Pernambuco, according to surveys conducted in 1997 and 2006. *Revista de saude publica* 2018; **52**: 48.

261 Vasconcelos MJOB, Rissin A, Figueiroa JN, Lira PICd, Batista Filho M. Diarrheal diseases and hospitalization of children under five years of age according to population-based surveys in the State of Pernambuco, Brazil, in the years 1997 and 2006. *Doencas diarreicas e hospitalizacoes em menores de cinco anos no estado de Pernambuco, Brasil, nos anos de 1997 e 2006* 2018; **23**(3): 715-22.

262 Vieira SCF, Gurgel RQ, Kirby A, et al. Acute diarrhoea in a community cohort of children who received an oral rotavirus vaccine in Northeast Brazil. *Memorias do Instituto Oswaldo Cruz* 2011; **106**(3): 330-4.

263 Volpicelli K, Buttenheim AM. Do Social Factors Predict Appropriate Treatment of Child Diarrheal Disease in Peru? *Maternal and child health journal* 2016; **20**(11): 2299-308.

264 Wang X-Y, Xu Z-Y, von Seidlein L, et al. Incidence of diarrhea caused by rotavirus infections in rural Zhengding, China: prospective, population-based surveillance. *The Journal of infectious diseases* 2005; **192**: S100-5.

265 Watts H, Gregson S, Saito S, Lopman B, Beasley M, Monasch R. Poorer health and nutritional outcomes in orphans and vulnerable young children not explained by greater exposure to extreme poverty in Zimbabwe. *Tropical medicine & international health : TM & IH* 2007; **12**(5): 584-93.

266 Webb AL, Stein AD, Ramakrishnan U, Hertzberg VS, Urizar M, Martorell R. A simple index to measure hygiene behaviours. *International journal of epidemiology* 2006; **35**(6): 1469-77.

267 Wierzba TF, El-Yazeed RA, Savarino SJ, et al. The interrelationship of malnutrition and diarrhea in a periurban area outside Alexandria, Egypt. *Journal of pediatric gastroenterology and nutrition* 2001; **32**(2): 189-96.

268 Wilson SE, Ouedraogo CT, Prince L, et al. Caregiver recognition of childhood diarrhea, care seeking behaviors and home treatment practices in rural Burkina Faso: a cross-sectional survey. *PloS one* 2012; **7**(3): e33273.

269 Woldemicael G. Diarrhoeal morbidity among young children in Eritrea: environmental and socioeconomic determinants. *Journal of health, population, and nutrition* 2001; **19**(2): 83-90.

270 Wu J, Yunus M, Streatfield PK, Emch M. Association of climate variability and childhood diarrhoeal disease in rural Bangladesh, 2000-2006. *Epidemiology and infection* 2014; **142**(9): 1859-68.

271 Wu J, van Geen A, e, et al. Increase in diarrheal disease associated with arsenic mitigation in Bangladesh. *PloS one* 2011; **6**(12): e29593.

272 Wu J, Yunus M, Streatfield PK, et al. Impact of tubewell access and tubewell depth on childhood diarrhea in Matlab, Bangladesh. *Environmental health : a global access science source* 2011; **10**: 109.

273 Xue J, Mhango Z, Hoffman IF, et al. Use of nutritional and water hygiene packages for diarrhoeal prevention among HIV-exposed infants in Lilongwe, Malawi: an evaluation of a pilot prevention of mother-to-child transmission post-natal care service. *Tropical medicine & international health : TM & IH* 2010; **15**(10): 1156-62.

274 Yassin K. Morbidity and risk factors of diarrheal diseases among under-five children in rural Upper Egypt. *Journal of tropical pediatrics* 2000; **46**(5): 282-7.

275 Yotthanooi S, Choonpradub C. A statistical method for estimating under-reported incidence rates with application to child diarrhea in Thai provinces bordering Cambodia. *The Southeast Asian journal of tropical medicine and public health* 2010; **41**(1): 203-14.

276 Yu JX, Zhu WP, Ye CC, et al. A cross-sectional study of acute diarrhea in Pudong, Shanghai, China: prevalence, risk factors, and healthcare-seeking practices. *Epidemiology and infection* 2017; **145**(13): 2735-44.

277 Zavaleta N, Kvistgaard AS, Graverholt G, et al. Efficacy of an MFGM-enriched complementary food in diarrhea, anemia, and micronutrient status in infants. *Journal of pediatric gastroenterology and nutrition* 2011; **53**(5): 561-8.

10. Appendix III: Measurements Contributed per Study in the Systematic Review and Meta-Analysis

Number of Measurements per Study	Number of Studies (%)
1	165(59.78%)
2	59(21.38%)
3	14(5.07%)
4	19(6.88%)
5	6(2.17%)
6	3(1.09%)
7	2(0.72%)
10	1(0.36%)
12	1(0.36%)
15	1(0.36%)
16	2(0.72%)
17	1(0.36%)
21	1(0.36%)
34	1(0.36%)

11. Appendix IV: CONSORT Checklist for Chapter 3

Section/Topic	Item No	Checklist item	Reported on page No
Title and abstract			
	1a	Identification as a randomised trial in the title	1
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	2-3
Introduction			
Background and objectives	2a	Scientific background and explanation of rationale	4-5
	2b	Specific objectives or hypotheses	5
Methods			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	6
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	12
Participants	4a	Eligibility criteria for participants	6
	4b	Settings and locations where the data were collected	6

Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	6-9
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	11
	6b	Any changes to trial outcomes after the trial commenced, with reasons	NA – no changes
Sample size	7a	How sample size was determined	11
	7b	When applicable, explanation of any interim analyses and stopping guidelines	NA
Randomisation:			
Sequence generation	8a	Method used to generate the random allocation sequence	6
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	6
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	6
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	6

Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how	7
	11b	If relevant, description of the similarity of interventions	NA
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	11
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	11
Results			
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	13
	13b	For each group, losses and exclusions after randomisation, together with reasons	13
Recruitment	14a	Dates defining the periods of recruitment and follow-up	13-15
	14b	Why the trial ended or was stopped	15
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	14
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	14
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	15-17

	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	NA
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	20
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	21
Discussion			
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	25
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	22-24
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	22-24
Other information			
Registration	23	Registration number and name of trial registry	27
Protocol	24	Where the full trial protocol can be accessed, if available	27
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	27-28

*I strongly recommend reading this statement in conjunction with the CONSORT 2010 Explanation and Elaboration for important clarifications on all the items. If relevant, I also recommend reading CONSORT extensions for cluster randomised trials, non-inferiority and equivalence trials, non-pharmacological treatments, herbal interventions, and pragmatic trials. Additional extensions are forthcoming: for those and for up to date references relevant to this checklist, see www.consort-statement.org.

12. Appendix V: Semi-Structured Interview Guide for Chapter 3

Introduction: Thank you for your interest in taking part in this interview. My name is Philemon Langat, and this is [NAME]. I work with UN-Habitat on the project using SMS messages to measure diarrhoea in Mwanza, which you may recall recently agreeing to be a part of. This project has been funded and is run by the University of Warwick in the UK. As part of the project, I hoped to interview 10 parents involved in the study about specific aspects of the diarrhoea survey, including for example, what worked well, what didn't work so well, what you found difficult or problematic etc. I want to speak to a range of people. You were selected to be one of the parents that I would (initially) like to interview. I will ask you a few questions – it will up to 30 minutes. You'll see me and my colleague [NAME] write down notes on what you say in order for us to write about the interview later. I will not record any information which could be used to identify you, and I will not audio or video record this interview. My discussion will either be reported in a summary, or with short quotes, ensuring that there is no identifying information. You will be able to find more information on this on the participant information sheet which I have handed you, and on your signed consent form. Are you happy for us to begin?

Questions	**Probes**
<u>Section 1: Daily Life, Overall Impressions and Use</u> First, I'd like to find out a little more about you and your family. Could you walk me through a typical day in your life, and a day in your household?	• What time do you wake up? • Do you work? What do you do? • Who do you live with? • Who cares for the children? • Who has what responsibilities in your home? • Who do you interact with on a daily basis? • How often do you use your mobile? And for what? • When do you go to sleep?

I'll get into a little more detail on this, but how did the SMS phone system fit into your daily life, and what were your overall impressions?	• Were they convenient? • Were they annoying? • Did you have to alter your life at all? • Did they come at a convenient time? • Was it good that you got reminders? • Did you change any behaviours based on this? ○ Perhaps hygiene behaviours; or paying more attention to child health/stools
<u>Section 2: Incentive and Airtime</u> You may have noticed that while I always sent you some airtime before the survey, TZS500, to make sure that you have enough airtime to complete the survey; I also occasionally sent you airtime after the survey, TZS1000, to say thank you. Did knowing that you would receive this airtime after the survey influence if you responded or not, and how you responded? Why?	• Did the incentive make it worthwhile to take part or not? • Did you feel pressured or not to give answers you think I'd want if I gave you incentive? • Did you talk to your friends or not about this at all? If so, what did you discuss? • Was TZS1000 a substantial incentive, or a small one? • Did caring or not about the incentive change over time?
<u>Section 3: Questioning Frequency</u> You may have also noticed that I adjusted how often I texted you: sometimes texting you every day, asking about diarrhoea over the past 24 hours; and sometimes texting you once in two weeks, asking about diarrhoea in the past two	• How was the frequency of the questioning – was it the right amount, or should it be more or less? • Did you feel that the questions were comfortable to answer, or where they too intrusive into your life? • When I asked you about what happened two weeks ago, rather than

weeks. Did you prefer one of these methods? And Why?	in the past day, were you able to remember? • Was receiving airtime a factor in you preferring frequent or infrequent messages?
<u>Section 4: Question Levels</u> As part of the survey, I sometimes asked a few more questions to carers reporting diarrhoea in their children – questions on blood in stool, vomiting, frequency of diarrhoea, and hospital visits. At other times, I did not. Did you have a preference to us asking extra questions or not?	• When I asked you several questions, did you find this to be intrusive? Or did you prefer to be able to share more with us? • Did I ever tell you to go to the hospital? If so, did you appreciate us doing this, and did you go?
<u>Section 5: Methods of Improvement</u> I're always looking for ways to improve my work, and greatly value your feedback. Would you be able to tell me a few of the things that you liked about the survey; things which made it more likely for you to take the survey; and things which you think I should introduce in the future? How about things that you did or did not like about the survey; things which made it easy or difficult for you; and things you think I should change? Would you be interested or not in taking part in a similar study in the future; and would you or would you not recommend a study like this to your friends?	• Did you feel comfortable answering the survey? Would you rather talk to a person, or do you prefer the privacy by phone? • Did your friends, family, and community pressure you to either complete or not complete the survey? • Was money a factor in completing or not completing the survey? • Were you able to use the technology easily, or was it confusing? • Could you easily understand the questions?

Conclusion Is there anything else which you could like to share with me?	

Thank you for taking part in this interview. I greatly appreciate your feedback, and will be using it in my future work. You'll continue to receive my text messages until August. Feel free to reach out to us if you have any questions. Remember that you can contact us on the information on the participant information sheet that I have given you.

13. Appendix VI: Questionnaire used in Cox's Bazar

Data Collection Log Book (Separated Data)

Name

Phone Number

Household Barcode SCAN BARCODE

Survey Data (Separate from Log Book)

RA | 1 | 2 | 3 | 4 |

EA - - - -

Date

Household SCAN

Barcode BARCODE

Resident Information (all respondents)

Name	Age	Sex	Education Level	Occupation	Health Conditions	Time Spent in Bangladesh	Relationship to Respondent
RESPONDANT							
SPOUSE IF PRESENT							
Child/Other Family							
Child/Other Family							
Child/Other Family							
Child/Other Family							
Child/Other Family							
Child/Other Family							
Child/Other Family							

Household Information (all respondents)

What is the main source of drinking water for your household?	Piped water into home	Piped water into plot	Piped water to tap	Borehole	Dug well	Spring	Rain	Truck	Cart	Surface Water	Bottled Water
Where is that water source located	In my home	In my plot	In a public space								
How long does it take you to get there?	Minutes										
Do you share this with other households?	Yes	No									
	If yes, how many	Households									

	If yes, how long do you have to wait for water	Minutes							
Do you do anything to make this water safe to drink?	Yes	No							
	If yes, what?	Boil	Bleach/Chlorine	Cloth Strain	Filter	Solar disinfection	Sit and Settle	Other:	
Can I see where you use the toilet?	Yes	No							
	If yes, describe:	Flush toilet to sewer	Pit Latrine	VIP Latrine	Flush toilet to Latrine	Composting Toilet	Bucket Toilet	Open Defecation	Other:
How many people use this toilet	People								

Is this toilet separated by sex	Yes	No				
Do you feel safe using this toilet during the day?	Yes	No				
Do you feel safe using this toilet when dark?	Yes	No				
How do you wash your hands?	Water alone	Water with Soap	Alcohol Based Sanitizer	Non-Alcohol Based Sanitizer	Does not wash hands	Other:
When do you wash your hands? (all that apply)	After Defecation	After cleaning a child's stool	Before Feeding a Child	Before Eating	Before preparing food	Other:
Do you own livestock?	Yes	No				

Where do you normally seek healthcare?										
OBSERVE MATERIAL OF FLOOR	Natural Floor	Wood	Bamboo	Tarp	Finished Floor	Other:				
OBSERVE MATERIAL OF ROOF	No roof	Thatch	Bamboo	Wood	Cardboard	Tin	Ceramic	Cement	Tent	Other:
OBSERVE WATER STORAGE VESSEL	Open Bucket	Closed bucket	Bottle	Jerry Can	Bucket	Other:				

Basic Survey (50% of respondants)

*IDENTIFY
OLDEST CHILD
UNDER 5*

Is this child breastfed?	Yes – Exclusively	Yes – partially	No
Has this child had 3 or more loose or watery stools any day in the past 2 weeks?	Yes	No	
Has this child had blood in their stool in the past 2 weeks?	Yes	No	
	If yes to either, did you seek treatment for either of these?	Yes	No
		If yes, where?	

[redacted]

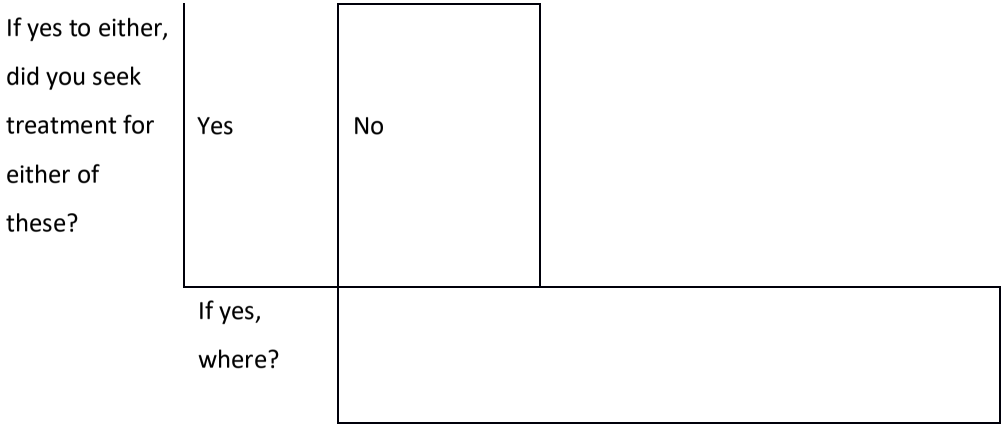
Enhanced Survey (50% of respondants)

IDENTIFY OLDEST CHILD UNDER 5

Is this child breastfed?	Yes – Exclusively	Yes – partially	No					
Mid-Upper Arm Circumference	cm							
*SHOW AMSTERDAM STOOL CHART*Has the child had any of these stools in the past 2 weeks	Consistency	A	B	C	D			
	Colour	1	2	3	4	5	6	
How long ago?	1	2	3	4	5	6	7	Other:
How many times on the worst day?	1	2	3	4	5	6	7	Other:

How many days did it last?	1	2	3	4	5	6	7	Other:
Did this child have fever in the past 2 weeks?	Yes	No						
Has this child had blood in their stool in the past 2 weeks?	Yes	No						
Did this child have vomiting in the past 2 weeks	Yes	No						
Has this child been eating properly in the past 2 weeks	Yes	No						
	If yes to either, did you seek treatment for either of these?	Yes	No					

		If yes, where?						
	If yes to either to these, did you provide treatment?	Yes	No					
		If yes, what?	Commercial ORS	Homemade ORS	Zinc	Traditional medicine	Antibiotics	Other:
Has this child had any unexplained rashes or redness in the past 2 weeks?	Yes	No						
Has this child had any unexplained pink eye or discharge from eyes in the past 2 weeks?	Yes	No						


If yes to either, did you seek treatment for either of these?
Yes
No
If yes, where?